Photocopy and Use Authorization

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at Idaho State University, I agree that the Library shall make it freely available for inspection. I further state that permission for extensive copying of my thesis for scholarly purposes may be granted by the Dean of the Graduate School, Dean of my academic division, or by the University Librarian. It is understood that any copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Signature _____

Date _____

TESTING ANCESTRY ESTIMATION METHOD COMPARABILITY:

OBSERVATIONS FROM A MANUFACTURED COLLECTION

AT ADA COUNTY CORONER'S OFFICE, BOISE, IDAHO

by

Hannah Brianne Dawson

A thesis

submitted in partial fulfillment

of the requirements for the degree of

Master of Science in the Department of Anthropology

Idaho State University

Summer 2019

To the Graduate Faculty:

The members of the committee appointed to examine the thesis of Hannah Brianne Dawson find it satisfactory and recommend that it be accepted.

_____

Samantha Blatt Ph.D.,

Major Advisor

_____

Katherine Reedy Ph.D.,

Committee Member

_____

Amy Michael Ph.D.,

Committee Member

_____

Roger Schmidt Ph.D.,

Graduate Faculty Representative

Thanks also goes to McKenzie Kennell who has been one of my greatest friends and supporters since I began my college journey in 2013. You've been an inspiration to me and have motivated me to work hard and stay strong, even in the face of adversity. You've been an absolute pillar in my life and I'm so lucky to call you sister.

Lastly, and most importantly, I must thank my family, who have been there from the beginning of all of my crazy dreams. Though you may not have always understood my passion and fascination with the subject, your support and confidence in my ability and future has been the ultimate driving force. You've taught me resilience, perseverance, patience, and a million other things that were all absolutely necessary to complete this project. Without your constant faith in me, I'd be nothing. You mean the world to me, and I love you all more than anything.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Ada County Coroner's Office .... ACCO

American Academy of Forensic Sciences .... AAFS

American Indian Female .... AF/AIF

American Indian Male .... AM/AIM

Amy R. Michael (for observer error) .... ARM

Anterior Nasal Spine .... ANS

Ascending Ramus Shape .... ARS

Author/Hannah B. Dawson .... HBD

Biorbital Breadth .... EKB

Black Female .... BF

Black Male .... BM

Bureau of Land Management .... BLM

Chin Shape .... CS

Common Era .... C.E.

Cranium and Mandible Only .... SO

Cranium and Postcranium .... CP

Cranium, Mandible and Postcrania .... SP

Cranium Only .... CO

Data Collection Procedures 2.0 .... *DCP 2.0*

Dental Arcade Width .... XDA

Discriminant Function Analysis .... DFA

Expectation-Maximization Algorithm .... EM

Testing Ancestry Estimation Method Comparability: Observations from a Manufactured

Collection at Ada County Coroner's Office, Boise, Idaho

Thesis Abstract – Idaho State University (2019)

Ancestry is an important component of the biological profile that estimates how an individual may have been perceived in life, and therefore aids in identification. Forensic anthropologists often combine both metric and nonmetric methods to increase overall accuracy. Few address the perceived comparability of these methods, however. This project compares ancestry estimations of individuals from the Ada County Coroner's Office in Boise, Idaho using several modern techniques and software programs. Metric and morphoscopic data were analyzed using FORDISC 3.1 (Jantz and Ousley 2010a), Optimized Summed Scoring Attributes (Hefner and Ousley 2014), Discriminant Function Analyses (Hefner 2015) and (hu)MANid (Berg and Kenyhercz 2017). All methods employed for this collection generated incomparable estimations. Discrepancies were likely due to the lack of standardization in reference collections and statistical techniques used for method creation. Continuing to build databanks of recent and diverse populations would elevate medico-legal study and allow further evaluation of multi-method analyses.

Key Words: ancestry estimation, forensic anthropology, coroner's office, manufactured collection, method comparability, FORDISC 3.1, OSSA scores, (hu)MANid, observer error, discriminant function analyses

Chapter I:

Introduction

1.1 Defining the Problem

Forensic anthropologists are tasked with identifying decomposing and skeletal human remains. The process of identification relies on estimations of biological parameters, known as the biological profile, which consists of providing an estimation for age, sex, stature, and ancestry of unknown individuals. Estimating ancestry is a key component of the biological profile as relating groups of phenotypic traits to biological and social groups can provide further evidence to assist law enforcement in the identification of an individual. Unlike other components to building a biological profile, estimation of ancestry has not been as rigorously refined due to constant changes spatially and temporally in the definition of race and its improper assumption that it is biological as well as cultural. Estimating ancestry from skeletal remains is difficult; historically, anthropology did not ascribe appropriate or scientifically sound meaning to ideas of "race". Reflecting on past qualms, the medico-legal nature of forensic anthropology nonetheless requires an estimation of ancestry that is understandable to those familiar with and applying sociocultural race terms, such as law enforcement.

Forensic anthropologists utilize both metric and non-metric (visual parameters) to distinguish ancestry. However, little work has been done to assess how forensic anthropologists apply and understand the contributions of these methods when used in combination. Due to the historical roots of race and typology of the skull in particular (the most useful element for assessment of ancestry [Betti et al. 2010]), it is necessary to understand the interplay of these methods. The statistical analyses that guide ancestry classification with expressed posterior

probabilities (i.e. the likelihood of an unknown individual being of the group predicted), and the data upon which the methods are based must also be assessed in respect to these methods.

1.2 Craniometric Analyses for Ancestry Estimation

A great deal of research focusing on the methodology of ancestry estimation has concerned craniometric data (Howells 1973, 1989; Jantz and Ousley 2010a; Plemons and Hefner 2016). The most widely popular metric method is the software program, FORDISC 3.1 (Jantz and Ousley 2010a). This program uses not only craniometric data, but also postcranial and mandibular metric data; however, more estimations of ancestry are possibly with the craniometric data as the cranium displays the most ancestral markers (Betti et al. 2010). FORDISC 3.1 has a large comparative database of measurements from positively identified individuals from forensic cases, documented skeletal collections, and historic collection in the United States and around the world. The program uses linear discriminant function analyses (DFA) (a statistical equation that assists in providing an estimation) from the Forensic Anthropology Data Bank (FDB) (University of Tennessee, Knoxville). The FDB utilizes craniometric data from 28 populations (Christensen et al. 2014; Howells 1973; Howells 1989), therefore expanding what is known as the traditional three-group (White, Black, and Native American) ancestry model. Thirty-six standardized measurements of the skull can be entered into the program, which then calculates the probability that the skull belongs to a particular ancestry group relative to all other groups considered in the analysis. This program, created in 1993, was revised in 1996 and again in 2005; the latter included new ways to analyze data and more global populations for reference (Jantz and Ousley 2010b:4).

1.3 Non-Metric Analyses for Ancestry Estimation

Forensic anthropologists also commonly use non-metric analyses of the midface, especially when certain cranial landmarks have been too damaged for measurements (Plemons and Hefner 2016:2). Non-metric observations visually record scores of traits that vary in frequency between different populations. A classic example of this would be a wide inter-orbital breadth being a common feature for American Whites; however, this is not exclusive to this group. It is important to realize that there is overlap in all non-metric trait expressions and that observer experience with these overlapping traits can produce biased results (Nakhaeizadeh 2013). Though most traits are found in some frequency in all populations, certain populations show the highest frequency of specific traits, which are then used to develop the probability of an individual belonging to that group. Such biased results are often created from using the *Gestalt* technique, in this case, the quick visual classification of sex or ancestry through years of conditioning and experience using such methods (Hefner 2009). Despite this, the use of morphometric or non-metric methods of recording discrete and nonlinear cranial traits, is common throughout biological anthropology (Hefner 2009; Hefner and Ousley 2006; Wheat 2009). Often, misclassification rates of non-metric traits are not assessed, so robust statistical evaluations of the classification of non-metrics are still being developed and improved (Klales and Kenyhercz 2014).

More recently, re-evaluation of qualitative traits for the estimation of ancestry has become important in forensic anthropology. This latest research involves documenting a continuum of macromorphoscopic traits of the cranium on an ordinal scale (e.g. Hefner and Linde 2018). Combined with statistical analyses and trait frequency distributions, macromorphoscopic studies can provide a useful alternative or supplement to craniometric

ancestry estimation. One data collection program called Osteoware includes a macromorphoscopic trait module, incorporating the work of Hefner (2009) and others (Hefner and Linde 2018), and is meant to act as a data collection tool for quantitative and qualitative observations. Morphometric methods have comparatively small databases, which may limit potential, but used in conjunction with metric analyses, may improve accuracy overall. Testing the reliability of the Osteoware program in the assessment of ancestry, Klales and Kenyhercz (2014) found that some macromorphoscopic traits are not representative of certain population groups (e.g. scores of 1 for anterior nasal spine and scores of 1 or 2 for inferior nasal aperture were not as frequent for Black populations in Klales and Kenyhercz [2014] as they were for Hefner [2009]). They also found that such qualitative traits could be analyzed statistically to increase repeatability and decrease bias. There has been much current re-evaluation regarding the subjectivity and standardization of non-metric methods (e.g. Hefner 2009; Plemons and Hefner 2016) where observer error and subjectivity did not appear to impact estimations greatly. However, there is no standard application by which to integrate metric and non-metric forms of data (if they can be integrated at all) or a true understanding of their ability to be applied synchronously. The gap in understanding the interaction of cranial metric and macromorphoscopic traits is the primary focus of this thesis.

1.4 Interaction of Craniometrics and Non-Metrics for Ancestry Estimation

Both metric and non-metric traits contribute to information about ancestry estimation, but the relationship and comparability between outcomes of methods that employ these data types has not been extensively explored. Both types of variables correlate with the size and shape of the cranium. Early literature assessing the interaction of metric and non-metric traits did not directly address forensic anthropology; many focused on primate crania and/or epigenetic

sitology of non-metric traits (Cheverud et al. 1979; Carpenter 1976; Corruccini 1974; Richtsmeier et al. 1984; Wilson 2010), but contributed to ancestry estimation methods. One of these studies found a positive correlation between craniometrics and non-metrics and associations between measurements of metric traits and non-metric traits in that they are all representations of adaptations to environmental pressures (Cheverud et al. 1979).

1.5 Research Questions

This project investigates how metric and macromorphoscopic cranial methods perform (statistically and practically) and compares the ancestry estimates generated in light of their varied statistical approaches and data sources. The present study contributes to the field of forensic anthropology in three main ways by: 1) contributing to existing literature exploring the performance and value of integrating macromorphoscopic and metric data by investigating if and how the two types of methodologies collaborate to provide similar estimations of ancestry; 2) test the limitations, misuse, and considerations of macromorphoscopic methods and software programs; and 3) providing new data to contribute to metric and morphoscopic forensic databanks.

Six methods of ancestry estimation were applied to a manufactured collection of skeletal remains in the Ada County Coroner's Office (ACCO) located in Boise, Idaho. Manufactured collections are defined by Komar and Grivas (2008) as "documented or forensic skeletal collections derived from donation or autopsy" (Komar and Grivas 2008:224). As such, they are not representative of the demography of the population as a whole.

This project addresses how these methods compare and what ancestries they generate. More importantly, this project also seeks to assess the reasons why the methods agree or disagree in their estimations, their limitations when used separately, and the value, misuse and

considerations of applying metric and morphometric observations of the skull to estimate ancestry.

Several questions with associated hypotheses are explored in this thesis:

1. Are metric and non-metric methods specifically (e.g. Berg 2015; Hefner 2015; Hefner and Ousley 2014; Jantz and Ousley 2010a) for estimating ancestry comparable (i.e. do they generate the same ancestry estimation)? How do methods with similar reference populations influence comparability?

    a. How do software programs (e.g. FORDISC 3.1 [Jantz and Ousley 2010a], and (hu)MANid [Berg and Kenyhercz 2017]) compare against one another?

2. If methods generate different estimations, what are the aspects of these methods that researchers need to consider? What improvements would elevate the accessibility and agreement between these methods? Can these methods be used to complement each other?

The main assumption is as follows: metric and non-metric methods will give similar ancestry estimations, particularly in cases of similar reference populations (that is, when ancestry estimation methods use the same or similar collections of given ancestries to refer to with their own data). To test, the methods applied to the ACCO collection were qualitatively compared against one another based on ancestry estimation outputs generated by each.

1.6 A Question of Representation in Manufactured Collections

Databanks and reference populations are built off data taken from what are normally termed "collections" (e.g. William M. Bass Donated Collection, Robert J. Terry Anatomical Skeletal Collection, Hamann-Todd Osteological Collection). These resources contain information from individuals who often willingly donated (not true of all collections) their

remains to the causes of these collections, though representation within them is impossible (see Chapter II). However, such collections that consist of remains left in coroners' or medical examiners' offices can be studied through different lenses with their different demographics. Remains here are left at the request of next-of-kin or may not be identified for this decision of donation to be made. Moreover, remains in this setting are likely to not be representative of the population that they derive from and have higher chances of having faced more uncommon trials in life than the majority of the overall population (e.g. poverty, substance abuse, structural violence, etc.) (Komar and Grivas 2008). Details regarding the representation of these obscure collections of remains will be discussed in Chapter II along with what can still be learned regarding groups like those observed at ACCO and larger collections as a whole. Though the ACCO collection is not representative, its information and data are still invaluable and useful for the scope of this research.

1.7 Chapter Outline

Chapter II discusses the separation of cultural race and biological ancestry along with reasoning for why ancestry estimation is justified in a forensic setting. The utility of software programs used to estimate ancestry and their significance in regards to manufactured populations is addressed.

Chapter III presents the materials and methods of this study. Descriptions of reference populations and statistical procedures of the methods used is given. Details of how the current data were recorded is included. Results of the methods are presented in Chapter IV with accompanying tables along with a discussion on observer error. The methods used are descriptively compared against each other within this chapter.

Interpretations and the conclusions of the results follow in Chapter V in which the research questions of this project are reviewed considering the data. The use and value of these methods in other contexts (e.g. bioarchaeological) is addressed. A discussion of the nature of the collection and limitations encountered in the study are given. Finally, a summary of the findings is presented along with potential future directions for the field of forensic anthropology.

Chapter II:

The Inter-Relationship between Cultural Race and Biological Ancestry

2.1 Introduction

Though forensic anthropologists acknowledge that there are no biological races, they do not deny that social races exist (Smedley and Smedley 2005). Governments and the public at large recognize these social races, as discussed by Smedley and Smedley (2005). Social races are described in the generic terms, "White", "Black", "Hispanic", etc. Furthermore, they possess social consequences, specifically regarding "privilege, power, and wealth" (Smedley and Smedley 2005:22). This ideology of social races results in consequences such as inequality and discrimination. In turn, these consequences only work to increase the disparity between groups, thus perpetuating the inaccurate concept of biological races (Sauer 1992; Wagner et al. 2017).

2.2 A Brief History of Cultural and Biological Race

The history of race is lengthy, dating back to the Romans with the Latin *generatio,* meaning "to beget" (Wolf 1992). As globalization became more pervasive, the definition and social consequences of "race" changed greatly. This could be seen through conquests (e.g. Vikings, Mongols, Romans, etc.) (Wolf 1992:3) and in specific locations as well, such as the debate and eventual war and emancipation from slavery in the United States through to the Civil Rights Movement.

Methods of differentiating populations have also been dynamic, moving from location, to spoken language, to religion, and to skin color or other phenotypic characteristics (Wolf 1992). Some of these conclusions were made after observing Non-Romans, composed of all those that the Romans came in to contact with. This included Germanic tribes, Vikings, Slavs, and even

groups further from their home such as the Huns and Mongols (Wolf 1992:3). It was assumed after contact with such groups that these peoples were "barbarians"; however, they were not so different that they could not become civilized to demonstrate an advancement in culture. Tribes of Germans, Vikings, Slavs and Saracens were most often considered to be the most redeemable. Other "barbarian" groups falling beyond a certain geographic range from western society appeared to have a form of civilization and therefore government, therefore still posing as a threat to the Roman empire as they challenged the Roman way of life. Mongols and Huns, along with others, fell into this category and were claimed "monstrous" (1992:3).

As Christianity began to influence the Roman empire, terms like "barbarians" and "monsters" were transformed to reflect the number of people capable of being converted to Christianity and therefore capable of being saved. Those who had previously been considered "barbarians" (e.g. Germans, Vikings, Slavs, etc.) were now "unredeemed" and those who were considered "monstrous" (Huns, Mongols, etc.) were catalogued as being "unredeemable" (Wolf 1992:3). During this period, behavior (not skin color) was what the populace judged, and this eventually turned into religious status, where Christianity was at the pinnacle of enlightenment and superiority. Those who fell into the "monstrous" or "unredeemable" categories were determined to be "natural slaves" as Aristotle called them (1992:2).

It is important to note that although characteristics such as religion, language, geographic location, and others may correlate to a particular phenotype (e.g. skin color), they do not represent biological traits that justify categorization into social races or ethnicities.

2.3 Early Attempts of Ancestry Estimation

In 1758, Carolus Linnaeus, the taxonomist responsible for the scientific classification still present today in biological studies, worked to make distinctions between peoples of different

geographic origins. Described by Stephen Jay Gould (1982) and Tishkoff et al. (2004), Linnaeus

endeavored to create four different races or subspecies of humans based off of Pliny the Elder,

whose works on "monstrous races" date from the first century C.E. (Johnson 2015:173-181).

Linnaeus' research recorded these four races as Europaeus, Asiatics, Afer and Americanus.

Others (see below) would term these groups as "Caucasian/White, Mongolian/Asian,

Ethiopian/Black, and American/Native American" (Gould 1982:3). Though focused on

classification, Linnaeus's belief that Whites were the superior race led to other scientists of time

building off of his foundational scheme.

This configuration would soon change from the four-race scheme proposed by Linnaeus

with the introduction of a fifth racial category, provided by Johann Friedrich Blumenbach.

Living from the 1750's to the 1840's, Blumenbach attempted to find races by observing variation

in physical characteristics present on the skull. Forensic anthropology frequently recognizes

Blumenbach for his erroneous and racist tendencies; however, Stephen Jay Gould (1982)

illuminates that Blumenbach was merely a product of his time (1982:4). Furthermore, Gould

believes that Blumenbach was even a progressive thinker for the period.

Blumenbach wanted to improve upon Linnaeus' ideas, leaving him to create a fifth

category known as "Malay" (1982:3), or the Australoid/Oceanic (Tishkoff et al. 2004:1). Gould

described Blumenbach's intentions as attempting to include more variation than what Linnaeus

had originally proposed (Gould 1982:3). While Linnaeus was much more overly racist by

modern standards, separating each of his four races into different subspecies of *Homo sapiens*,

Blumenbach recognized that all populations eventually bled together to form one cohesive

species (1982:5). Stephen Molnar makes this point at length *Human Variation: Races, Types,

and Ethnic Groups* (2006). However, with the introduction of a fifth category, a pyramidal

scheme placed one race at the pinnacle, above all others. Blumenbach found the "Caucasoid" group to be the most aesthetically beautiful; this contention cemented this group's position as the most superior among scientists and the public (Gould 1982:7-8). Other scientists have employed Blumenbach's views on subjective beauty (albeit potentially misinterpreting them) many times thereafter to explain their contention that Caucasians were the better of the races and that other races were degenerative of this original, Caucasoid form. Many anthropologists perpetuated this principle; however, there were others that remained uncomfortable with the idea of assigning race to any population.

In the nineteenth century, Samuel G. Morton examined skulls from North America, with particular focus on how to differentiate between races. In his book, *Crania Americana* (1839), Morton attempted to show that some races scientifically were better than others by using the cranial capacity of skulls as evidence. Concerns of his potentially biased methodology were addressed by Gould in *The Mismeasure of Man* (1980); however, some have found Gould's work to also be biased in that Gould looked for errors in Morton's work (Rushton 1997). Nevertheless, Morton's position within the American school of anthropological thought at the time was important and did lead many within North American to believe that those of the Caucasian race were superior to other races due to a larger cranial capacity.

Among those who also agreed with the idea of some races being superior to others was Earnest A. Hooton who studied human evolution and would come to focus on cranial morphoscopic traits and their relation to criminology and race (Garn and Giles 1995). Some of his work (Hooton 1939) was criticized as being similar to Cesare Lombroso, where certain cranial traits and shapes were assumed by Hooton to be the markings of criminals. Much of the morphoscopic features that Hooton assessed during his time eventually became a part of what

would be called the "Harvard List." These traits are still in use today (and in this thesis) for many morphoscopic methods (Hefner 2018). However, these traits were not standardized until Finnegan and McGuire (1979) – more than 20 years after Hooton's death. Nevertheless, Hooton's mark on anthropology in terms of craniometrics and morphological analyses was greatly important and is still crucial in today's medico-legal setting.

Franz Boas and his students would eventually play a fundamental role in the destruction of this biological race concept within anthropology. In *Science*'s 1931 edition, Boas discussed concerns he had about the conflation of physical and psychological characteristics in order to describe certain population groups. His argument "Race and Progress" (1931) stated that identifying biological race in modern populations was essentially impossible, given the amount of "intermating" (Boas 1931:3). Many of his statements fit very well in modern understandings and opinions of race, as he also stated that functionally in a biological sense, there was very little difference between populations. Furthermore, he made it clear that the environment and living conditions greatly impact not only functional adaptations, but also how an individual mentally handles such circumstances (1931:4-5).

Discussion would move beyond Boas in the early twentieth century, however, with discussions regarding biological determinism (Caspari 2003) in the 1960s. This concept is defined by the thought that social race (which was determined at birth) was correlated to a certain level of evolutionary history for the species (Caspari 2003:67). Individuals would not be able to amount to anything more than what was already allowed for that particular race. Boas did argue against this concept; however, given that it was not deeply intertwined with the construction of racial typologies, it endured. Only when the idea of races being equivalent to

subspecies was dismantled by evolutionary biologists (e.g. Lewontin 1972 [see below]) was the idea of biological determinism largely disregarded by physical anthropologists.

2.4 Globalization and Genetics

Within the academic community, it is more commonly accepted that race is not a biological reality, but working this awareness into the public mindset has had limited success. Finding a solution to this problem is challenging as there is not one distinct definition of race that can be argued against. Not only is it defined different temporally and spatially, but it is *felt* differently as well, being a systemic viewpoint that continues to persist (Smedley and Smedley 2005). Specific focus to this issue is commonly applied to the status of human rights in the current political climate. Figures in political office can assist by promoting equality, though even this approach may not always be enough. Other organizations promoting equal human rights through campaigns also are capable of improving the matter. Textbooks used may have the potential at the college level (and perhaps with the ability to extend into high school classrooms) to help decrease the popularity of the scientific race concept and explain that biologically, there is no basis for races existing at all. Anthropologists of all subfields are also arguably obligated to work to dismantle the stigma attached to certain social races (Wagner et al. 2016:326).

Biological anthropologists are aware now that variation is far greater within populations than between them (Lewontin 1972; Molnar 2006) and increased globalization only continues to contribute to homogenization of the species. Globalization remains a heady topic, involving more than just the admixture of different cultures. This concept was further discussed as a whole by Lewontin (1972, 2006), claiming that because of increased human mobility, admixed individuals are more prevalent than they have ever been before (Lewontin 1972). This trend brought about its own assortment of challenges for the forensic anthropologist, as individuals of

mixed social race are given more of an opportunity to decide for themselves to choose the social race with which they identify. Though this challenge certainly has and had forensic implications (see section 2.5), Lewontin (2006) focused more on the medical consequences.

Lewontin states that when medical professionals ask of a patient's medical history, they are not asking how a person identifies themselves, rather where their ancestors originated from as there are conditions that appear more prevalent with some populations than in others (Lewontin 2006). This claim, however, may still prove somewhat inflammatory as differences in medical conditions (e.g. sickle cell anemia or diabetes) between population groups may suggest a significant difference in genetics. This specificity for such conditions could thereby support the position that there must be substantial genetic differences between these populations to warrant a distinction of biological races. This claim, Lewontin argued, still proved irrelevant and incorrect in the overall discussion. He stated that genetic variation within populations is, as a whole, much greater than that between populations. In fact, 85% of total genetic variation comes from within any given population simply due to the aforementioned globalization process and "intermixing" process as Boas (1931) termed it (Lewontin, 1972).

Though there do seem to be sets of linked genes (genes inherited together more often than would be expected due to chance) that work together to present similar phenotypes in closely related populations, it still must be understood that these phenotypes juxtaposed alongside every other phenotype show that changes in appearance are more gradually distributed along what are called clines or gradations than they are definitely and absolutely distributed (Bulatao and Anderson 2004; Molnar 2006). Furthermore, these genetic differences that result in these stereotypical phenotypic differences (e.g. skin color, hair form, nose shape) only amount to approximately 6-10% of DNA (Lewontin, 2006:2).

Geneticists and anthropologists, like Lewontin (1972, 2006) and Gould (1982), recognized that no significant part within the genetic code exemplified differences between individuals who came from diverse ancestral backgrounds. Populations from everywhere in the world were in one way or another related and exhibited their differences in clines across a spectrum where each cline blends into another without notable boundaries. When an individual is viewed singularly, it would seem to be relatively straightforward to estimate his or her social race. However, when individuals are put in larger groups, it becomes clear that one cannot find quantifiable boundaries of where one biological "race" ends and the other is meant to begin.

2.5 Some Concerns and Roles of the Forensic Anthropologist

As it becomes apparent that there are no such things as biological races, questions arise of how forensic anthropologists are able to estimate ancestry when race is not visible on the human skeleton. Like the race concept, ancestry estimation also has its more recent history of methods, most of which are rooted more heavily in science than in conjecture (albeit flawed science in some instances) (Christensen et al. 2014, Sauer et al. 2016). Many methods are discussed in Sauer et al.'s (2016) chapter regarding ancestry estimation within Blau and Ubelaker's text, *Handbook of Forensic Anthropology and Archaeology* (2016). A considerable number of these methods involve working on past populations that are not temporally significant due to secular changes or trends over generations defined as "the process that results in a change in mean size or shape of a population from generation to the next" (Bogin 1999:243-244). Despite this, it is worth noting that they may still be highly applicable to historic and bioarchaeological cases. Through these secular trends, populations as a whole are concluded to be dynamic and adaptable, further distorting any potential of clear-cut lines demarcating boundaries of races. These and

other similar issues make creating and maintaining modern and representative skeletal collections challenging, if not impossible.

2.5.1 Limitations with Reference Collections/Databanks

One of the primary features within forensic anthropology are reference skeletal collections. There are many roles that these collections fulfill in academia. While not entirely intended to demonstrate human variation and be representative of certain populations (e.g. local, regional, etc.), many anthropologists believe that they accomplish this task of representation as Komar and Grivas discuss (2008). This belief of representation is not new, as Wood et al. (1992) address this issue regarding bioarchaeological samples and disease in regards to paleodemography. Their idea (called the Osteological Paradox) that the visible dead in the archaeological record did not represent the population they died in was sensational and is recognized with great value in biological anthropology today. Komar and Grivas (2008) address representation in a more modern sense with forensic anthropologists currently relying on skeletal reference collections and databanks to facilitate positive identifications.

One frequently employed metric analysis technique is used with the computer program known as FORDISC 3.1 (Jantz and Ousley, 2010a). This program employs a discriminant function analysis (DFA) to classify individuals into ancestry groups based on the collective similarity of reference craniometric data within the Forensic Data Bank (FDB). It also has data collected by Howells (1973, 1989) that attempt to demonstrate historic global populations. This database was created in 1986 with the goal to provide extensive information regarding skeletons dating from a contemporary period (University of Tennessee, Knoxville).

Through the use of metric and non-metric analyses, many databanks for software programs reference these collections. Specific to this study, the author defines collections to be

groups of skeletonized individuals with known or unknown details of their biological profiles (i.e. sex, age, ancestry and stature). Databanks are available for forensic anthropologists who are not able to have contact or measure the collections directly, making them invaluable resources for many.

The representativeness of these collections, however, is lacking, as Komar and Grivas (2008) clearly define that the demographics within collections are often drastically different from the surrounding, living demography. Some groups of individuals (e.g. elderly, White males) donate their remains to collections more often than any other group (Komar and Grivas 2008:224) which potentially translates to these individuals living at a higher rate of income, resulting in less skeletally stressed individuals given their status (DeWitte and Stojanowski 2015:414).

For all variables observed – sex, age, ancestry, cause and manner of death – the demographic within these collections was not representative of the living population of New Mexico, where the study was conducted (Komar and Grivas 2008). Komar and Grivas (2008) conclude that regardless of the size of the collection or databank (e.g. FORDISC 3.1 uses the FDB which totals over 2,400 documented cases (even these may be suspect with fluctuating definitions over ancestry over time) with biological profile information [University of Tennessee, Knoxville]), it *is not* and *cannot* be representative of the population even at a local level.

Though the size of a reference collection does not increase its representativeness, by adding more individuals to databanks, forensic anthropologists still capture a greater breadth of variation within a population group. Increasing variation in databanks creates greater chances of providing an ancestry estimation that matches that of an unknown decedent's self-identification.

Many forensic anthropologists from around the United States still work to continuously add to the FDB from the cases that they work, though reference sample sizes for some populations (e.g. American Indian/Native American males and females) continue to remain low. Likewise, in 1996, FORDISC added data collected from Howells' craniometrics; however, the methods used by W.W. Howells in his 1973 book, *Cranial Variations in Man* (1973) generate results that warrant caution (see below).

Within forensic anthropology, FORDISC 3.1 continues to be the most commonly used method of ancestry estimation and, as such, sample sizes of several reference populations remain a concern, largely due to size and representation. This is especially true of historic and prehistoric groups as well as contemporary Native Americans. The National Conference of State Legislatures federally recognizes four Native American tribes existing in Idaho alone (Coeur d'Alene Tribe, Kootenai Tribe, Nez Perce Tribe, and Shoshone-Bannock Tribe) (National Conference of State Legislatures) and such groups are likely even more challenging to properly account for through these databanks.

Within the FDB, there are only 32 American-Indian Females and 59 American-Indian Males (labeled Ameri-Indian Females/AF and Ameri-Indian Males/AM in the software respectively) (Jantz and Ousley 2010b:15). Given the variation among Native Americans (e.g. Shoshone-Bannock of the Intermountain West vs. Hopi of the American Southwest) and how remains become a part of these skeletal collections (frequently without tribal consent), adequate representation within each population is highly unlikely.

The FDB is an important tool, and with its accessibility through FORDISC 3.1 and over 2,400 documented cases entered into the databank (Jantz and Ousley 2010b:19), it is widely used (Williams et al. 2005:340-341). However, without greater discussion or acknowledgement of the

limitations of sample sizes and reference populations, there are concerns regarding how reliable and useful such a tool and database may be, and as such the FDB should be used with caution, as argued by other forensic anthropologists (2005:341). It is still undecided as to how this problem will be resolved and some individuals (Albanese and Saunders 2006; Komar and Grivas 2008) continue to question how little phenotypic variation is represented, and questions of how to make variation more present in collections are still common. Despite this lack of a united response to this problem, software like FORDISC 3.1 and others try to expand their recognition of phenotypic variation by introducing more samples to their databanks.

Increasing the sample sizes of reference populations, however, does not resolve the fact that outputs from FORDISC 3.1 are specific to one possible ancestry (though all groups included in the output should be considered by the anthropologist), not recognizing that every individual is admixed to some degree (Christensen et al. 2014:227). Furthermore, the software, while capable of notifying the user of outliers, cannot recognize the impact that taphonomy and pathology may have on a set of remains. As such, it is the responsibility of the user to ascertain which measurements are viable. Given this challenge and others such as skeletal idiosyncrasies and secular change, the observer must consider all components before acknowledging them as relevant and recording a measurement or non-metric trait for any set of methods. (Hefner and Linde 2018:4; Christensen 2014:233, 301, 314-315; Boyd and Boyd 2011:1408).

Within forensic anthropology, many assume that the individual set of remains will be from a modern population, allowing forensic anthropologists to use the FDB as their reference group, despite not being representative. With a suspected bioarchaeological sample, the utility of the FDB is less certain as secular trends dictate that populations even of the same ancestry have the capacity to change over time, as was seen in the discussion of Boas mentioned above (Boas

1912, 1931; Bogin 1999). Databanks recognize secular changes specific to time periods, though their relation to modern secular changes becomes more ambiguous. Secular changes in this respect become smaller in magnitude as the temporal boundaries of the databank become larger.

2.5.2 Howells' Contribution

W.W. Howells addressed the challenge of variation in bioarchaeological samples, though did not address secular change beyond American Whites and American Blacks in the 19th and 20th centuries. He obtained measurements on bioarchaeological specimens starting in the 1970's (Howells 1973). He made his measurements available to other researchers at the time, allowing them to employ the same statistical techniques he used on ancient skulls from around the world. Such opportunities allowed anthropologists like Richard Jantz and Douglas Owsley (2001) to observe Native American populations from approximately 4,500 BP to beyond even 7,000 BP (Jantz et al. 2001:146) and record the various tribes that were on the North American continent at that time. These individuals are more similar to Eastern Asian populations described by Howells and less similar to today's population of Native Americans (Jantz and Owsley 2001).

Given the popularity of FORDISC, its creators incorporated Howells' measurements into the program. Not all of the populations Howells observed are included, though there are populations from every continent as options to include. Along with these are additional American categories like 19th century White and Black males, along with 20th century White and Black males (Jantz and Ousley 2010a).

However, despite these advancements with Howells' measurements and application of multivariate analyses, the observer should recognize the limitations with the Howells dataset and its applicability in assessing ancestry for groups with small sample sizes and sex estimations provided through FORDISC 3.1 (Elliott and Collard 2009; Messer 2013; Ousley et al. 2018;

Tasa and Vogel 2016). It is argued by some that the methodology employed by Howells to estimate sex involved him choosing skulls that appeared to best represent not only the sex, but the population ancestry as well (Elliott and Collard 2009; Messer 2013). Furthermore, the dataset created by Howells and used in FORDISC 3.1 does not have temporal notes for a large part of his sample (excluding 19th and 20th Century Whites and Blacks).

Others have found that using the Howells dataset, especially in cases where the individual may be of Native American/American Indian ancestry along with other groups (e.g. Nubians and Egyptians), FORDISC 3.1 does only slightly better than chance at predicting such an ancestry (Tasa and Vogel 2016; Ousley et al. 2018:72). Tasa and Vogel (2016) document the case of Kennewick Man, an individual from the Pacific Northwest, whose Native classification with Howells' dataset was tenuous and likely due to chance with other groups from Howells employed as well. They recommend that results generated by not only Howells but also the FDB be interpreted with caution. They advise the researcher regard to the sample sizes of populations used, or the occasional inability for results from the Howells tab in FORDISC 3.1 to correspond to results produced from samples in the FDB (Elliott and Collard 2009; Stephens 2000).

2.5.3 Destructive Methods

Concerns regarding the value and representation of reference collections for forensic and bioarchaeological specimens are certainly apparent within forensic anthropology; however, there are other methods not specific to metrics (e.g. FORDISC 3.1) used that come with their own limitations or apprehensions.

In cases where more opportunities for observation beyond metrics or morphoscopics are available, genetic testing may also be an option to a forensic anthropologist. This testing requires destructive methods, however, and may require more funds and time than the forensic

anthropologist is appropriated. Nevertheless, the results garnered in such analyses may be of importance and may indicate different ancestry than what morphoscopic or craniometric methods generate and may more readily provide a positive identification.

2.5.4 Non-metric Analyses of Ancestry Estimation

Given a lack of full understanding of all possible impacts of secular trends on ancestry (both the impacts known and unknown), many anthropologists placed their focus on morphological or non-metric aspects of the skeleton that were correlated to certain ancestral populations yet did not emphasize temporal significance. Many of the morphoscopic traits that are observed and recorded are variable in appearance. Therefore, an individual will most likely present ancestral markers that were historically attributed to separate groups. Hefner and Ousley (2014) document Optimized Summed Scoring Attributes (OSSA) as an example of this variation (see Methods Chapter III). These attributes are identified through scoring several specific landmark traits (e.g. anterior nasal spine, inferior nasal aperture) using diagrams. These scores generally will rank anywhere from zero to five depending on the feature observed and will later be converted to a binary number and then summed to result in a potential ancestry.

While the OSSA method is commonly used (Christensen et al. 2014; Sauer et al. 2016), there are still improvements to be made to this method, as it only allows for an individual to be characterized as American Black or American White. Despite this difficulty, the technique does come with an understanding that individuals compose pieces of a gradation/cline and are not necessary members of a distinct group, as multiple traits are observed.

Gradations in any degree are likely depending on the method employed (e.g. OSSA); as a result, complications of too many traits to consider are possible and need to be limited to a smaller number. Some forensic anthropologists choose to combat this complexity by employing

metric methods for ancestry estimations as well. This practice allows a higher degree of *probability*, but it never indicates a higher degree of *certainty* (Sauer et al. 2016)

2.5.5 The Ethical Role of the Forensic Anthropologist

Despite these available tools, along with others not discussed here, many (e.g. Williams et al. 2005) ask whether or not providing law enforcement with an ancestry estimation is ethical as it potentially enables law enforcement to think that cultural race (what group a person identifies with) and biological ancestry are the same. It may also wrongly indicate to law enforcement that that there are such things as biological races.

Others, like Sauer (1992) and Sauer et al. (2016) argue that providing ancestry estimations to law enforcement falls within the realm of duties of the forensic anthropologist. While law enforcement may not always understand that an ancestry given by a forensic anthropologist is an *estimate* and is always a *true reflection* of biological variation, the information is often necessary for the investigative process to move forward. This estimate becomes an important piece of culturally relevant data, as it can be *correlated* but not *absolutely equal* to the social race that the unknown individual may have identified as in life or how surrounding persons (e.g. friends, family, or strangers) would have identified them. When law enforcement is attempting to identify missing persons from remains, a forensic anthropologist is responsible for providing as full and complete of a report as possible. This is done in order to identify the decedent more efficiently for law enforcement as well as families that are grieving and waiting for the identification of a loved one.

2.6 Summary

Though estimating ancestry takes practice and observer error is always possible, ancestry estimation is still capable of contributing to the identification of unknown human remains. Its

relation to race is limited, as ancestry relies more heavily on scientific observations to demonstrate the relative frequencies of skeletal features in particular ancestral geographic populations. These populations may have a race culturally ascribed to them, but such an attribute falls outside of the realm of the forensic anthropologist and falls more to the responsibility of law enforcement.

Chapter III:

Materials and Methods

3.1 Introduction

This chapter outlines the collections and methods used for this project and the utility and

limitations of both. The history of the sample used in this thesis, including how they were

obtained and came to be currently housed, will be presented. Finally, the metric, morphoscopic,

statistical, and comparative methods used to estimate the ancestry of these individuals in the Ada

County Coroner's Office will be described.

As discussed in Chapter I, ancestry estimation methods were applied to each type of

completion category (see section 3.2.2) with the intent to see how the methods compare in their

results. The ancestry estimation methods discussed below were used in the hope that they would

corroborate one another as such data could be submitted to databanks with a higher degree of

probability that an individual culturally identified to the population group estimated. When

different approaches (i.e. macromorphoscopic and craniometrics) result in the same ancestry

estimation, the general assumption is that the estimation is strengthened (Plemons and Hefner

2016:2). As stated in Chapter II, however, that increased probability does not necessarily indicate

increased accuracy in the correlation between self-identity and skeletal ancestral markers (Sauer

et al. 2016).

3.2 Materials

3.2.1 The Ada County Coroner's Office (ACCO) Collection

Cranial and postcranial metrics and non-metric cranial data were collected from 71

skeletonized individuals curated at the Ada County Coroner's Office (ACCO), located in Boise,

Idaho. The collection consists of individuals who died from the 1940s through the 1980s. Only

49 (N=49) individuals were used in the current study mainly for preservation reasons that may have obscured craniometric or non-metric trait landmarks. Some individuals in the collection originated as unclaimed remains, while the majority of the remains were donated to the ACCO by other county coroners in Idaho and the greater Pacific Northwest.

The remains of 16 of the total 71 individuals was transferred to the ACCO by the Department of Geology from Idaho State University in Pocatello, Idaho in 2015 (Laura Larson, personal communication, 2019). The late Dr. David E. Fortsch, a member of the Department of Geology at Idaho State University served as Idaho's vertebrate osteologist and was frequently contracted to consult in cases involving the recovery and analyses of human remains. He coordinated with law enforcement throughout the state of Idaho in the 1980s and 1990s in order to provide biological profiles for these unidentified individuals (Laura Larson, personal communication, 2019).

Before curation at ACCO, these remains were retained in Fortsch's personal collection. After his death, the collection was transferred to ACCO. At the time of acquiring and arranging transport of this collection, many of the remains were discovered by ACCO to be commingled. ACCO research staff spent three years sorting, cleaning and documenting the commingled remains into distinct individuals. In the summer of 2018, the author (HBD) was recruited to estimate the biological profiles for the individuals in this collection and the other individuals present at the ACCO. Of the 16 individuals, only 10 met the criteria necessary to be included in this study (see next section).

3.2.2 The Sample

Given variation in preservation of all remains (including the Fortsch Collection), seven preservation categories were generated. Names of these categories, their definitions, numbers

within each category, and methods applied to them are presented in Table 3.1. As the cranium

demonstrates the highest degree of morphological and metric variation associated with ancestry

of any skeletal element (Betti et al. 2010), methods that focused on the cranium were used.

Given the importance of the skull in regard to ancestry estimation (Betti et al. 2010), the observer

was unable to utilize many ancestry estimation techniques on cases within the ACCO that did not

include a skull.

Individuals were omitted from the study sample if ancestry could not be effectively

evaluated using the remains present. Exclusions were made for three reasons: (1) the remains

were too fragmentary for ancestry estimation, (2) only postcranial elements were present, or (3)

the individual was a subadult (estimated as less than 15 years of age). Until puberty, individuals

do not exhibit traits found in adults that mark ancestry, making sex and ancestry difficult to

estimate (Christensen et al. 2014:236). Individuals included in the study were aged 15 years or

older, as this is the earliest at which the facial region of the skull is likely to attain maturity

(Bastir et al. 2006:644). In cases where only the cranium and/or mandible were present, the

eruption of the third molars from the gums marked an age estimate of 17+ years (American

Dental Association 2006). Ubelaker (1989) marks that the third molars erupt from the alveolar

bone between the ages of 15 and 21; as such, age ranges provided by the American Dental

Association (2006) were used, despite the difference between gum and alveolar bone eruption.

After the individuals were evaluated based on age (i.e. 15+ years of age) and

preservation, a total of 49 (N=49) individuals remained and were available for the analyses in

this study. Other methods specific to aging adults were also applied (see section 3.3.2).

All metric and inventory analyses were recorded on hard-copy paper data forms. These

were scanned and stored electronically in the ACCO database. A GPM Model 106 spreading

caliper and the Neiko Tools Model 01407A Electronic Digital Caliper sliding calipers were used to take measurements from the cranium and postcranium. Ancestry estimation techniques based on dental morphology had the potential of being assessed; however, upon observation, only one individual out of the entire sample retained more than half of their dental arcade that did not include either severe cases of dental wear, or signs of invasive dental care (e.g. fillings). As such, dental morphology methods are not examined here.

Table 3.1. Categories, Abbreviations, Sample Sizes and Methods Applied for Each

Category	Abbreviation	n	DFA with Cranial Traits	FORDISC 3.1	(hu)MANid	OSSA
Cranium Only	CO	20	x	x		x
Cranium and Mandible Only	SO	11	x	x	x	x
Cranium, Mandible and Postcrania	SP	6	x	x	x	x
Mandible Only	MO	11		x	x	
Postcrania Only[a]	PO[a]	1[a]		x		

[a]Singular individual (IND16.2) included based on likelihood that this individual and IND16.1 are the same individual.

3.3 Sex and Age Methods

While ancestry estimation methods were the focus of this study, sex and age were estimated as well and were done so before ancestry estimation techniques were applied. These components (sex and age) can have impacts on how ancestry estimations are interpreted and will be discussed in greater detail below.

3.3.1 Sex Estimation

Estimations of sex and age were done through the observation of morphoscopic traits scored as present/absent or as one of a group of ordinal or nominal categories (Christensen et al. 2014:57). Both sex and age estimations were given prior to estimation of ancestry. Sex estimations have been stated in the past to be useful in aiding ancestry estimations, particularly in regard to FORDISC as they reduce "the effects of size differences" between the sexes (Christensen et al. 2014:233). However, as the majority of cases in the ACCO are not positively

identified individuals, ancestry analyses using a program (e.g. FORDISC 3.1, [hu]MANid) were used with sexes pooled as an incorrect sex estimation may impact how a program estimates ancestry (Messer and Getz, in press).

In cases where postcranial elements were absent or unable to be linked to other skeletal elements and only a cranium was represented of an individual (n=35), methods used for sex estimation relied on recording cranial morphoscopic or non-metric traits (Acsádi and Nemeskéri 1970; Buikstra and Ubelaker 1994). These traits include the nuchal crest, mastoid process, supra-orbital margin, supra-orbital ridge/glabella and the mental eminence and were taken from Buikstra and Ubelaker (1994).

When present, traits observed and scored from the postcrania included the ventral arc, subpubic concavity, ischiopubic ramus (Phenice 1969), greater sciatic notch (Buikstra and Ubelaker 1994), and preauricular sulcus of the innominates (Milner 1992). All traits listed here are provided in diagrams in Ubelaker and Buikstra (1994). Six individuals were not complete enough to estimate sex through morphoscopic means. The sex distribution of the sample is listed in Table 3.2. Individuals in this table are also marked with their status of preservation (Cranium Only, Skull Only, etc.). IND16.1 and IND16.2 were marked as two individuals in ACCO documents; however, it is likely that these remains are from a single individual. As this is not confirmed, they were kept as separate individuals and were not analyzed through any method as one. However, IND16.2's ancestry estimation for the postcrania is included in Chapter IV alongside IND16.1 for comparison purposes.

IND39's cranium had undergone taphonomic damage, resulting in only the neurocranium being preserved. As such, dentition was not available to estimate age. Though IND39's cranial

sutures were open, the extensive development of the cranial bones and overall size is consistent

with the individual being an adult. IND39 is listed as a special age estimation case in Table 3.2.

Table 3.2. Sample (N=49) with Age and Sex Estimations and Degree of Completion.

Case Number	Status of Preservation	Age (years)	Estimated Sex (morphoscopic)
IND1	CO	17+	Male
IND2	CO	17+	Female
IND3	CO	Indeterminate	Male
IND4	CO	17+	Female
IND5	CO	Indeterminate	Female
IND6	SP	35-44	Male
IND7	CO	17+	Female
IND8	SO	17+	Female
IND9	CO	Indeterminate	Female
IND10	SO	17+	Male
IND11	SO	17+	Male
IND12	CO	Indeterminate	Male
IND13	CO	17+	Male
IND14	CO	17+	Male
IND15	CO	Indeterminate	Male
IND16.1[a]	SP	45-50	Male
IND16.2[a]	PO	Indeterminate	Indeterminate
IND23	SP	15-20	Male
IND25	MO	17+	Female
IND26	MO	17+	Male
IND27	MO	17+	Male
IND28	SO	17+	Female
IND29	SO	17-21	Male
IND30	SO	17+	Male
IND31	MO	17+	Male
IND32	MO	17+	Male
IND33	MO	17+	Male
IND34	SO	17+	Indeterminate
IND35	SO	17+	Indeterminate
IND36	SO	17+	Indeterminate
IND37	SO	17+	Male
IND38	CO	17+	Female
IND39[b]	CO	17+	Female
IND40.1	MO	17+	Female
IND40.2	MO	17+	Male
IND45	CO	17+	Male
IND46	CO	17+	Male
IND47	SP	17+	Female
IND48	MO	17+	Male
IND49	MO	15-21	Female
IND50.1	CO	17+	Male
IND50.2	MO	17+	Indeterminate
IND51	CO	17+	Female
IND52	CO	17+	Indeterminate
IND53	CO	17+	Male
IND54	SP	17+	Male
IND55	SP	30-39	Male
IND56	SO	17+	Male
IND57	CO	17+	Male

[a]Likely the same individual

[b]Special age estimation case

3.3.2 Age Estimation

Both morphoscopic observations of bone and dental eruption were used to estimate age (American Dental Association 2006; Brooks and Suchey 1990; Buikstra and Ubelaker 1994; Isçan 1984, 1985; Lovejoy et al. 1985; Ubelaker 1989a, 1989b). For individuals with all permanent dentition erupted and/or total epiphyseal fusion of long bones, observation of the development and degenerative topography of pubic symphyses (Brooks and Suchey 1990), sternal rib ends (Isçan 1984, 1985), and iliac auricular surface (Lovejoy et al. 1985) were used to estimate age. Destructive methods (e.g. histology) were not performed for an age estimation and are outside the scope of this thesis and research questions.

3.4 Metric Analysis for Ancestry Estimation

3.4.1 FORDISC 3.1

Twenty-four standard craniometric measurements based on Buikstra and Ubelaker (1994) were taken. These measurements are listed in Appendix A. All metrics preserved were collected; however, not all measurements were used in every analysis due to preservation. These data were analyzed with FORDISC 3.1 using reference populations from the FDB (Jantz and Ousley 2010a).

The use of the Howells dataset in FORDISC 3.1 (Jantz and Ousley 2010a) was excluded for reasons discussed in Chapter II: small sample sizes of certain groups, questionable reliability for sex estimations done by Howells (1989), inapplicability for Native American/American Indian samples in the Pacific Northwest, lack of temporal information for many, and its documented occasional disagreement with the FDB (Elliott and Collard 2009; Messer 2013; Stephens 2000; Tasa and Vogel 2016). Given also that the other software to be employed, (hu)MANid, (see section 3.4.2) required the use of composite groups that had modern

individuals included, it was decided that only the modern databanks would be compared for ancestry.

Many forensic anthropologists use the Howells tab when there would be a strong indication of an individual being from an historic context (i.e. before the 20th century [Jantz and Ousley 2010b:45]). Identifying historic cases was important specifically for the ACCO and such a distinction was asked to be provided for their records primarily to determine forensic significance and if an investigative case needed to be opened. A common indication of remains being historic included extensive untreated dental pathologies (i.e. periapical osteitis or significant dental caries). Other indicators included the majority of teeth displaying a high degree of dental wear in cases where the teeth may have been used as tools in life, thus potentially distinguishing them from modern peoples. These features are all potentially indicative of little to no modern dental care (Rösing et al. 2007:82).

All measurements taken were entered into FORDISC 3.1. Both males and females for all groups were initially included in each analysis per the recommendation of Messer and Getz (In press) and Stepwise Forward Wilks was employed in each case. Using Stepwise Forward Wilks aids in selecting the fewest number of measurements that collectively provide the greatest differentiation between groups while avoiding overfitting using highly correlated measurements. This procedure also allows for reference groups with small samples to be included in the analysis.

While it is recommended by the Jantz and Ousley (2010b) that sample sizes be at least three times the size of the number of measurements taken, all groups were used in first runs of the program. The FORDISC 3.1 Help Version 1.48 File acknowledges that with the FDB's sample sizes for population groups, the risk of overfitting is now reduced and not likely a

concern (Jantz and Ousley 2010b:83). Cases where the population sample size was not at least three times the size of the number of measurements taken were still noted.

With each program run, population groups with F typicalities less than the chosen alpha level of 0.05 were removed from analysis unless specified. Typicalities (in FORDISC 3.1 there are F, Chi, and R typicalities) assess how *typical* the measurements inputted into the program are for the predicted group. If these are less than 0.05, this is an indication that the measurements inputted are not typical or common to the group, despite what a posterior probability may indicate (Jantz and Ousley 2010b). The posterior probability is simply the probability of the unknown individual being of the predicted group in respect to the other reference populations used for the run; these will always add to one, whereas typicalities are independent of other reference populations utilized in an analysis. Measurements that were listed as being four standard deviations away from the group mean were removed from the analysis. This was repeated for all inter-/intra-observers and their measurements.

When more than one element was present (i.e. cranium, mandible, postcranial elements), each element was run singularly through the software for analysis to see how ancestry estimations would differ between elements and the reference groups available for each. This would simulate cases where one or more element was missing upon recovery in a forensic or bioarchaeological context. Postcranial elements were run together under the postcranial tab. This was done due to the variance in reference sample sizes for other elements (i.e. there are more reference populations used for the cranium and the mandible than for the postcranium) and to ensure that each element would be analyzed with as many reference groups as possible. In the case of the mandible, maximum ramus height (XRH) was not used in order to include the Guatemalan Male sample since XRH data is not available for that sample.

FORDISC 3.1 uses discriminant function analyses (DFA), specifically linear discriminant analyses (LDA) in order to produce its classifications and also forces a set of measurements into a group after discriminant scores and Mahalanobis' Distances (D^2) are calculated. LDA also only uses one group centroid/mean, thereby assuming only one multivariate normal distribution for a group. As Kenyhercz and Berg (2018) describe, LDA assumes that each unknown individual being analyzed through FORDISC 3.1 derives from one mutually exclusive group (Kenyhercz and Berg 2018:36).

3.4.2 Ancestry Estimation from the Mandible

Cases that included the mandible in any state of preservation were also subjected to specific morphoscopic and metric examinations using the method developed by Berg (2015). The method aims to estimate sex and ancestry based on reference samples from many global populations. The reference samples for this particular method derive from many different global populations. These are outlined with the populations they are meant to represent and their respective sample sizes in Table 3.3. These collections have been supplemented since the time of original publication (2015); however, Berg and Kenyhercz (2017) have created a free, Web-based graphical user interface (GUI) to input metric and morphoscopic data and generate sex and ancestry estimates. The GUI's (hu)MANid databank continues to grow by accepting metric and morphoscopic data from forensic anthropologists. Sample sizes used for this study are current as of April, 2019 and are listed below in Table 3.3, which also includes the abbreviation of each population. "M" demarcates a male sample whereas "F" demarcates a female sample.

Table 3.3. Descriptions of Composite/Pooled Population Groups in (hu)MANid GUI (Adapted from Berg and Kenyhercz 2017).

Composite Sample	Populations Represented	Composite Abbrev.	Details of Sample	Total Size
American Black	19th Century American Black, 20th Century American Black	BF, BM	78 F; 113 M	191
American White	19th Century American White, 20th Century American White	WF, WM	145 F; 296 M	441
American Indian	Arikara, Hohokam	AIF, AIM	44 F; 65 M	109
Pooled Hispanic	Guatemalan, Hispanic	HISPF, HISPM	14 F; 119 M	133
Northeast Asian	Chinese, Korean	NEAF, NEAM	15 F; 199 M	218
Southeast Asian	Cambodian, Thai, Vietnamese	SEAF, SEAM	88 F; 315 M	403

The GUI allows the user to run metrics, morphoscopics, and morphometroscopics (the combining of metric and morphoscopic data) with a linear discriminant analysis (LDA) as well as a mixture discriminant analysis (MDA). Linear discriminant analyses (LDA) are similar to principal component analyses (PCA) in that they both recognize variation within classes (i.e. ancestry population groups). LDA, however, looks to maximize separation between these groups by maximizing differences between group means (i.e. group centroids) while PCA uses variation as a means of prediction. Unlike LDA, a mixture discriminant analysis (MDA) goes beyond one normal distribution of data for a population group. Instead, MDA uses a mixture of normal/Gaussian distributions and uses an Expectation-Maximization (EM) algorithm to continuously adjust group means through a k-means cluster in order to achieve maximized separation between groups (Kenyhercz and Berg 2018:36).

Kenyhercz and Berg (2018) predict that the use of MDA will be more accurate in estimations than applying just an LDA to the data. As LDA uses only one group centroid, LDA may have a tendency to wrongfully force an unknown individual into a classification through the calculation of the Mahalanobis' Distance (D^2) which takes a distance in respect to three different dimensional planes, while MDA uses a Euclidean Distance instead which takes a distance

between two points in two dimensions. By creating more than one group centroid, MDA group assignment is more likely to be accurate (2018:36).

The program allows for a Stepwise Forward Wilks function to be employed; however, classification accuracy using this function does not appear to improve over when the function is not used at all (Berg and Kenyhercz 2017:1595). As such, Stepwise was not employed in any case when estimating ancestry using the GUI.

Berg and Kenyhercz (2017) note that composite groups (i.e. the combination of specific geographic and temporal groups into one broad group) and individual groups (e.g. Thai females versus the composite Southeastern Asian Females) likely should not be run together (Berg and Kenyhercz 2017:1596). This is likely due to the overlap generated when using composite and individual groups together, as some groups are comprised of exactly the same individuals (e.g. the composite Hispanic Female group consists entirely from the individual group Guatemalan Female group). Given the specificity of the individual groups and the unlikelihood of an unknown individual from ACCO deriving from them (e.g. Cambodian group), only composite groups were utilized. This also allowed for more comparability for the reference groups that FORDISC 3.1 (Jantz and Ousley 2010a) employs through the FDB.

Though only composite groups were used for analyses, the size of the sample groups was also important. The author (HBD) chose only groups that had a sample size of at least three times the size of the number of measurements/morphoscopic traits observed to run for analysis. This decision was based on previous instruction for multiple regression analysis, primarily as discussed in the FORDISC 3.1 Help Version 1.48 File (Jantz and Ousley 2010b:83). Though this eliminated several groups possible for a result (e.g. HISPF and NEAF), it also eliminated the risk of overfitting the data to conform to these groups in some cases. While HISPF and NEAF have

small sample sizes (n=14 and n=15 respectively), they both consist of only Guatemalan Females and modern Korean Females only. As it seems unlikely that either of these populations would be strongly represented in the ACCO, the elimination of these groups from analysis when sample sizes were not large enough for the number of measurements/traits was deemed worth the risk.

Both sex and ancestry estimations are generated through the use of multivariate discriminant function analyses. Measurements followed the standards outlined by Berg (2015), despite possible contradicting measurement techniques outlined in Buikstra and Ubelaker (1994). Berg (2015) created two measurements (mandibular body breadth at M2/M3 junction [TML23] and dental arcade width [XDA]) specifically for of this method. One measurement (mandibular body breadth [TML]) was redefined due to the field largely taking the measurement inconsistently or incorrectly with previous published material instructions (Moore-Jansen et al. 1994). Macromorphoscopic traits and their abbreviations with descriptions are outlined in Table 3.4 below. Following these in Table 3.5 are the instructions of measurement techniques for mandibular metrics that are not outlined or that differ in instruction from *Standards* (Buikstra and Ubelaker 1994).

Table 3.4. Abbreviations/Descriptions of Macromorphoscopic Traits of Mandible (Adapted from Berg 2015).

Morphoscopic Trait (with abbrev.)	Scoring System	Description/How to Observe (Berg 2015)
Chin Shape (CS)	Blunt; Pointed; Square; Bilobate	Viewed from above (superiorly)
Lower Border of Mandible (LBM)	Straight; Undulating; Partial Rocker; Rocker	Score this trait by placing the mandible on a flat surface
Ascending Ramus Shape (ARS)	Pinched; Wide	Viewed from lateral aspect
Gonial Angle Flare (GAF)	Inverted; Straight; Slight; Medium; Everted	Viewed from above (superiorly)
Mandibular Torus (MT)	Present; Absent	The mandibular torus is a bony protuberance of varying size and shape on the lingual surface, below the alveolar margin, typically in the region of the premolars
Posterior Ramus Edge Inversion (PREI)	Straight; Slight; Medium; Turned	This trait is observed on the posterior one-third of the ascending ramus

Table 3.5. Mandibular Measurements with Instructions Specific to Method (Adapted from Berg 2015).

Measurements (with abbrev.)	Tool Recommended	Descriptions/How to Measure (Berg 2015)
Mandibular Body Breadth at Mental Foramen (TML)	Sliding Calipers	Maximum width of the mandibular body taken at the mental foramen. The measurement is taken from the superior view to inferior direction and the caliper arm should be parallel to the flat surface on which the mandible is resting.
Mandibular Body Breadth at the M2/M3 Junction (TML23)	Sliding Calipers	Maximum mediolateral breath of the corpus taken at the level of the articulation between second and third molars. The sliding caliper arm should be parallel to the surface the mandible is resting on. The measurement location usually corresponds to a medial-lateral thickening of the mandible at that location.
Dental Arcade Width (XDA)	Sliding Calipers	Maximum breadth of the dental arcade at the level of the posterior-most points of the third molar sockets on the lingual surface. If necessary, a line should be drawn perpendicular to the ramus body and the tooth crypt to mark the measurement locations. If the third molars are congenitally absent, the measurement should not be taken.

Metrics. All mandibles were first analyzed using only metrics within the GUI. Data from each mandible was run twice, the first through a linear discriminant analysis (LDA) and the second through a mixture discriminant analysis (MDA).

Population groups were eliminated from each of the runs one at a time based on Chi-squared typicality values. If a value fell under the alpha level of 0.05, then this group would be removed from the run. Not until all groups possessed a Chi-squared typicality of equal to or greater than 0.05 would the author report the predicted group, the distance from the centroid, the Chi-squared typicality, the cross-validation and other assessments. The same process was repeated for all metrics taken by the inter-observers and again with the intra-observer data to ensure a control over how the estimations and analyses would be interpreted.

Morphoscopics. Each of the six possible morphoscopic traits was converted to a numerical data type as designated by Berg and Kenyhercz (2017) on the GUI. All mandible morphoscopic data was run using both LDA and MDA and data would be collected in the same method as it was for metrics. Mandibles that only had three observable morphoscopic traits (n=2) were not able to be run using the software. As with metrics, only composite reference groups were used as reference populations and sample sizes of each population needed to be at least

three times the size of the number of morphoscopic features observed to be included. Inter-/intra-observer data was not collected for any morphoscopic data as there was limited variation in skill level when assessing morphoscopic traits as opposed to large variation when taking measurements.

Morphometroscopics. The GUI allows for both metric data and morphoscopic data to be used together, allowing more aspects of skeletal variation to be captured. Both LDA and MDA were run with all data and the same criteria as described above with reports generated after all requirements for typicalities were met. As morphometroscopic data requires an input of morphoscopic features, inter-observer and intra-observer data were not collected in this instance.

3.5 Non-Metric Analyses for Ancestry Estimation

Below are descriptions of each of the non-metric ancestry estimation methods employed in this study. Explanations of how each was employed along with a description of the statistics used within each method are provided as well. All results generated from these methods are compared to the results generated by metric methods descriptively in Chapter IV to determine whether or not these methods provide similar or equal ancestry estimations.

3.5.1 Optimized Summed Scoring Attributes (OSSA)

Optimized Summed Scoring Attributes (OSSA) (Hefner and Ousley 2014) uses cranial morphoscopic traits. These are defined as visible, morphological traits that are observed and generate ordinal or categorical data. Traits used by Hefner and Ousley (2014), Hefner (2015), Hefner and Linde (2018) are macromorphoscopic traits on the skull and are scored with a range anywhere between 0-5 or on a presence/absence basis depending on the specific trait (Christensen et al. 2014:57).

Hefner and Ousley's method (2014) entails observing and scoring six macromorphoscopic attributes (see Table 3.6 for definitions and instructions for observation): anterior nasal spine (ANS), inferior nasal aperture (INA), interorbital breadth (IOB), nasal aperture width (NAW), nasal bone contour (NBC), and post-bregmatic depression (PBD). Each trait was scored according to the trait definitions and diagrams.

An OSSA score was generated using the scoring spreadsheet presented in Hefner and Linde (2018:318). Observations were then converted to a binary score (0 or 1). The scores were added together to form a summed compressed score ranging from 0-6. Individuals scored from 0-4 were estimated by this method to be American Black, while scores from 4-6 were estimated to be American White (2018:289). A score of 4 was indeterminate, though may possibly indicate an ancestry estimation of American White given a higher frequency of American Whites with this score as opposed to American Blacks in the authors' sample when developing the method (Hefner and Linde 2018:289; Hefner and Ousley 2014). As such, individuals with this score of 4 were marked as most likely being American White.

While this method has been cited in several textbooks on forensic anthropology and has become popular (Christensen 2014:231; Sauer et al. 2016; Hefner 2018), the reference populations from which it is derived are limited to American Black and American White individuals. However, the original OSSA method without additional groups was used as a fundamental starting point for the observation of macromorphoscopics on the cranium and is still widely used today (Hefner 2018).

3.5.2 Discriminant Functions for Two- and Three-Group Analyses

Hefner (2015) subsequently expanded the sample from American Black and American White to include those of Hispanic ancestry along with a seventh non-metric trait, nasal

overgrowth (NO). This method abandons converting morphoscopic traits into a binary. Instead, Hefner provides a set of discriminant function equations using either a three-way discriminant function or a two-way function. The three-way function allows for an estimation of Hispanic to be possible while the two-way is particular to American Blacks and American Whites. Both the three-way and the two-way provide the option of a seven-trait, five-trait and three-trait equation to be used depending on how many morphoscopic scores were collected. Discriminant function scores resulting from these equations were then compared to given centroids for each population group. In the case of a three-way discriminant analysis, if the score did not definitively suggest the first population group (American Black) then a secondary equation was calculated to distinguish between Hispanics and American Whites.

Cross-validation for all provided equations suggests that when the unknown individual is identified (e.g. by self, kin, etc.) as being of American Black, American White, or Hispanic ancestry, the likelihood of the method correctly assigning the unknown to that group is greater than assigning the ancestry on chance alone.

Again, it must be acknowledged that between Hispanics, American Whites, and American Blacks, the sample present within the ACCO is likely not representative of the local or regional population in and around Idaho. Nevertheless, this method was employed to document not only the changes in ancestry estimation techniques, but to also recognize how the additional group may have affected the predicted ancestry.

3.5.3 Expansion of Macromorphoscopic Traits

Hefner and Linde (2018) extended the trait list to include 17 total macromorphoscopic attributes; 11 had been previously used by Hefner (2009). This expansion of traits was done for a number of reasons, primarily to show a greater breadth of human variation than previous

methods that only used six or seven morphoscopic traits. It also further demonstrates a single

trait will not result in a single predicted ancestry. The entire list of 17 traits used also for OSSA

and Hefner's (2015) prior seven non-metric traits are presented in Table 3.6. Instructions for how

to score these traits can be found in Table 3.7 and are taken from Hefner and Linde (2018).

Scores were collected; however, given the lack of a usable discriminant function analysis,

software, or program, these trait scores were not used to create an ancestry estimation in this

thesis. However, these scores may be used to corroborate other estimations upon further

development of the technique to eventually become a method (and therefore are available in

Appendix Table A-1).

Table 3.6. Definitions and Scoring Systems of Macromorphoscopic Cranial Traits (Hefner and Linde 2018).

Macromorphoscopic Trait (with abbrev.)	Scoring System	Feature Definition (Hefner and Linde 2018)
Anterior Nasal Spine (ANS)	1.0-3.0	Small bony feature located at the inferior border of the nasal aperture
Inferior Nasal Aperture (INA)	1.0-5.0	Transition from the nasal floor to the vertical portion of the maxilla
Interorbital Breadth (IOB)	1.0-3.0	Space between both orbits relative to the facial skeleton
Malar Tubercle (MT)	0.0-3.0	Caudally protruding tubercle on the inferior margin of the maxilla and zygomatic bones
Nasal Aperture Shape (NAS)	1.0-3.0	Lateral contours of the nasal aperture and their greatest projection on the lateral margins
Nasal Aperture Width (NAW)	1.0-3.0	Width of the nasal aperture relative to the entire facial skeleton
Nasal Bone Contour (NBC)	0.0-4.0	Curvature of the midfacial region, specifically the nasal bones and frontal processes of both maxillae at 1 cm below the cranial landmark nasion
Nasal Bone Shape (NBS)	1.0-4.0	Relative contour of the nasal bones at their lateral edges
Nasal Overgrowth (NO)	0.0-1.0	Projection of the lateral border of the nasal bones at their inferior edge beyond the maxilla at the cranial landmark nasale inferious
Nasofrontal Suture (NFS)	1.0-4.0	Shape of the suture separating the nasal bones from the frontal bones
Orbital Shape (OBS)	1.0-3.0	Shape of the eye orbits
Postbregmatic Depression (PBD)	0.0-1.0	Depression along the sagittal suture posterior to the cranial landmark, bregma that is not a result of pathology
Posterior Zygomatic Tubercle (PZT)	0.0-3.0	Posterior projection of the zygomatic bone as viewed laterally on the cranium; also referred to as the marginal process, is viewed at approximately the midorbit level
Supranasal Suture (SPS)	0.0-2.0	Secondary complex suture superior to the cranial landmark nasion that may persist into adulthood; may also be referred to as supernasalis
Transverse Palatine Suture (TPS)	1.0-4.0	Shape of the course of the transverse palatine suture on the hard bony palate
Palate Shape (PS)	1.0-4.0	Contour of the dental arcade, defined by the curvature of the hard palate in the transverse plan, as viewed from the occlusal/inferior aspect of the maxillae
Zygomaticomaxillary Suture Course (ZS)	0.0-2.0	Suture between the maxilla and zygomatic

Table 3.7 How to Score Macromorphoscopic Cranial Traits (Hefner and Linde 2018).

Macromorphoscopic Trait (with abbrev.)	How to Score (Hefner and Linde 2018)
Anterior Nasal Spine (ANS)	The cranium should be viewed laterally to assess the degree of projection of ANS; do not score if the individual is edentulous
Inferior Nasal Aperture (INA)	The left portion of the inferior nasal aperture should be assessed; ignore subnasal grooves in the nasal floor
Interorbital Breadth (IOB)	Is assessed by using a ratio of IOB to the overall facial breadth. IOB is essentially visually observing the space between two cranial landmarks at left and right dacryon
Malar Tubercle (MT)	The side with the greatest expression of MT should be scored. This trait is assessed by placing a transparent ruler at a point approximately 0.5 cm lateral to the inferior terminus of the zygomaticomaxillary suture, extending to the deepest superior incurvature on the maxilla. A score is assigned on regarding the extent of the bone protruding past the ruler's inferior edge. Observations should not consider any tubercles on the lateral portion of the zygomatic arch
Nasal Aperture Shape (NAS)	This trait is scored by placing the cranium in anatomical position and examining the anterior portion of the face. The relative shape of the nasal aperture is assessed, with

Macromorphoscopic Trait (with abbrev.)	How to Score (Hefner and Linde 2018)
	emphasis in determining the location of the greatest projection of the lateral margins of the nasal aperture lateral walls
Nasal Aperture Width (NAW)	NAW should be assessed using a ratio of nasal aperture width to the overall width of the facial skeleton
Nasal Bone Contour (NBC)	Visual assessment of NBC is not recommended due to a high level of inter- and intraobserver error. It is recommended that this trait be scored with a contour gauge, as this tool permits a more objective assessment of NBC. To score this trait, the user should place the contour gauge directly on the nasal bones 1 cm below nasion. Keeping the contour gauge perpendicular to a transverse plane and parallel to the sagittal place, the observer should apply gentle but firm, consistent pressure until the deepest points on the contour gauge are reached. This process should be repeated multiple times along the length of the contour gauge to ensure correct assessment of NBC. Each needle on the contour gauge represent 1 mm. Make sure the needles on the gauge do not separate when pressing the instrument on the nasal bones
Nasal Bone Shape (NBS)	This trait is visualized and assessed while holding the cranium in its approximate anatomical position, with the anterior view directly in front of the observer. In order to make a correct observation, keep in mind the position of the nasal pinch (if present) and the amount of lateral bulging (if present). Do not consider the frontonasal suture, the nasal suture, or the symmetry of the nasal bones
Nasal Overgrowth	This trait is visualized by close inspection of the inferior lateral border of the left nasal bone where it articulates with the maxilla. It may be useful to gently run your finger along the border of the maxilla and nasal bones at nasale inferious to determine whether a projection is present. If the left side is damaged, the right side may be used to assess the trait. If both nasal bones are missing, fractures (ante- or peri-mortem), or damaged, do not score nasal overgrowth.
Nasofrontal Suture (NS)	To score the trait, hold the cranium in anatomical position, with the anterior portion of the cranium facing the observer. Do not assess this trait from a lateral view. When scoring, disregard the symmetry of the nasal bones and score for overall shape of the suture. If the nasal bone exhibit extreme pinching of the superior border (as in NBS score 4), do not score NFS
Orbital Shape (OBS)	To score the trait, hold the cranium in anatomical position, with the anterior portion of the cranium facing the observer. Do not assess this trait from a lateral view. When scoring, disregard the symmetry of the nasal bones and score for overall shape of the suture. If the nasal bone exhibit extreme pinching of the superior border (as in NBS score 4), do not score NFS
Postbregmatic Depression (PBD)	To score this trait, hold the cranium in a lateral profile view and look for a depression posterior to bregma. It may be helpful to palpate the area. PBD can be assessed using the contour gauge to detect small expressions of this trait. Be careful scoring individuals with obliterated sutures because the bone along the suture sites may be elevated, giving the appearance of a depression for the surrounding bone.
Posterior Zygomatic Tubercle (PZT)	This trait is scored by viewing the cranium in lateral view. By placing a transparent ruler on the frontal process of the zygomatic that extends from the two cranial landmarks, frontomalare posterale to jugale, the degree of protrusion of the tubercle can be assessed. A score is assigned on the extent of the bone protruding past the ruler's edge
Supranasal Suture (SPS)	This trait is scored by viewing the cranium anteriorly
Transverse Palatine Suture (TPS)	To assess this trait, view the cranium inferiorly, at the hard palate. Follow the TPS and note how deviation occurs near the intersection with the median palatine suture. If the suture is obliterated, do not score TPS. Do not consider slight undulations of the suture when scoring, particularly the area directly adjacent to the median palatine suture
Palate Shape (PS)	To assess PS, hold the cranium from the inferior view and focus on the occlusal/inferior aspect of the maxilla. Draw an imaginary line through the midline-projection of the individual teeth along their mesio-distal axis. Do not score edentululous individuals as resorption can cause the palate shape to change
Zygomaticomaxillary Suture Course (ZS)	This trait is scored from viewing the cranium in the anterior view. The left side is scored. Any infraorbital sutures should be ignored when scoring ZS

This technique, while still in development, gathers ordinal data from submissions made

by other forensic anthropologists and will eventually input them into a growing databank. The

Macromorphoscopic Databank (MaMD) databank consists of data from over 7,500 individuals from more than 20 populations (Hefner 2018; Hefner and Linde 2018:4). These individuals are derived from eight different collections: University of Tennessee, Knoxville, TN: William M. Bass Donated Skeletal Collection; Cleveland Museum of Natural History: Hamann-Todd Collection; Texas State University, San Marcos, TX: Donated Skeletal Collection; Khon Kaen University, Kohn Kaen, Thailand; Michigan State University, East Lansing, MI; National Museum of Natural History, Smithsonian Institution, Washington D.C. (2018:xxxi). This databank is meant to assist in forensic casework; it may also have applicability in prehistoric instances and can be used by bioarchaeologists with focuses in the United States (2018:4).

3.6 Inter-/Intra-observer Error

In addition to the author, measurements were taken by three observers with three experience levels to assess inter-observer error. These levels were classified by the author as novice, intermediate (where the observer includes herself prior to this study), and expert. The main observer/author determined the level of proficiency and experience based upon how much previous experience any given observer had with taking measurements. In order to be recognized as a novice, the observer had experience measuring fewer than 10 sets of remains in total. To be classified as intermediate, the observer had experience measuring greater than 10 but no more than 20 sets of remains. An expert observer had experience measuring more than 20 sets of remains.

Observers were not required to examine morphoscopic features as there are noted challenges with doing this when examining observer error (Hefner 2009; Byrnes et al. 2017). There were no varying levels of experience between observers for morphoscopics (e.g. no

observer had any experience with mandibular morphoscopic traits), resulting in the author placing focus onto only metric features.

It was chosen arbitrarily by the author that measurement agreement would be a 2 mm or less discrepancy between the inter-/intra-observer and the original measurement. This determination of agreement would be for both cranial and mandibular measurements. Any problematic measurements (i.e. measurements that had greater than a 2 mm discrepancy between observers) were compared against measurements that had been labeled as potentially problematic in the literature (see Chapter V).

Of the 49 individuals observed in this study, 23 (~47%) were measured to assess inter-observer error. Half of those measured (n=11) were measured by Joseph Effingham (JHE), undergraduate (novice level) and Shanda L. Putnam, B.A. (SLP) (intermediate level). Both individuals identify anthropology as their primary program of study and both have experience in forensic contexts. Skeletal remains measured by these observers had previously been estimated by the main observer (HBD) as being of an ancestry other than Native American.

The other half of the sample (n=12) were measured by Dr. Amy R. Michael, PhD (ARM) (expert level) and included all individuals estimated by the main observer to potentially be Native American in ancestry. Table 3.8 contains information on which individuals were subjected to inter-observer and intra-observer error calculations. The observers' initials are listed alongside these individuals along with the author's initials (HBD).

Table 3.8. Individuals Subjected to Observer Error. JHE=Joseph H. Effingham; SLP=Shanda L. Putnam, B.A.; ARM=Amy R. Michael, PhD; HBD=Hannah B. Dawson, B.A.

Case Number	Inter-observer(s)	Intra-observer
IND1	JHE; SLP	HBD
IND2	ARM	HBD
IND3	ARM	HBD
IND4	ARM	HBD
IND6	JHE; SLP	HBD
IND11	ARM	HBD
IND12	JHE; SLP	HBD
IND13	ARM	HBD
IND14	JHE; SLP	HBD
IND15	JHE; SLP	HBD
IND27	JHE; SLP	HBD
IND32	JHE; SLP	N/A
IND33	JHE; SLP	HBD
IND37	ARM	N/A
IND38	N/A	HBD
IND39	ARM	N/A
IND40.1	JHE; SLP	HBD
IND40.2	JHE; SLP	HBD
IND46	ARM	HBD
IND47	ARM	HBD
IND50.1	N/A	HBD
IND50.2	N/A	HBD
IND51	JHE; SLP	HBD
IND52	ARM	HBD
IND53	ARM	HBD
IND54	N/A	N/A
IND57	ARM	HBD

The measurements done by Dr. Michael were done at the request of the Ada County Coroner's Forensic Supervisor, Laura Larson, to assess the likelihood of these individuals being Native American. If these individuals were estimated as Native American, the Bureau of Land Management (BLM) and the State Historic Preservation Office (SHPO) of Idaho were contacted during the summer of 2018. This was done to commence repatriation in federal compliance with the Native Americans Graves Protection and Repatriation Act of 1990 (NAGPRA).

The individuals measured for inter-observer error (n=23) were then re-measured by the main observer approximately two months after initial measurements to assess intra-observer error. Some individuals (IND32, IND37, IND39 and IND54) were mistakenly not remeasured by the author. Each observer's exact measurements were run through the same forms of analyses as other metric data originally gathered by the observer and respective ancestry estimations would be provided (see Chapter IV).

3.7 Summary

In total, this thesis used two metric methods (Jantz and Ousley 2010a; Berg 2015; Berg and Kenyhercz 2017) and applied them to the cranium, mandible, and postcranium when applicable. Four non-metric methods were also used (Berg 2015; Berg and Kenyhercz 2017; Hefner 2015; Hefner and Ousley 2014) for the cranium and the mandible. Two software programs were utilized in total (FORDISC 3.1 and [hu]MANid). These programs were applied to the cranium (FORDISC 3.1) and the mandible (FORDISC 3.1 and [hu]MANid). Morphoscopic scores taken from the mandible were also analyzed using (hu)MANid. Additional macromorphoscopic trait data (Hefner and Linde 2018) was recorded (see Appendix A) but ancestry was not estimated using these features, as a method is not yet available.

Observer error was assessed in terms of variation in ancestry estimation outputs by the software programs using strictly metric data.

All of these methods were applied when applicable to the ACCO collection and their individual results are provided in the following chapter (Chapter IV). A comparison of all methods and their predicted ancestries is also provided. Discussion on observer error significance is provided in the following chapter also.

Chapter IV: Results

4.1 Introduction

This chapter presents the results of the ACCO dataset. Presented first are the results from methods described in Chapter III and software outputs with a comparison of these results provided at the end of the chapter. While an agreement of the utilized methods could suggest that the unknown individual self-identified as the overall predicted ancestry, the author found a large disagreement between the results generated by these methods. This demonstrates a need for forensic anthropologists to find more cohesive ways to create and apply methods in both medico-legal and bioarchaeological settings.

4.2 Metric Analysis for Ancestry Estimation

     4.2.1 FORDISC 3.1

Table 4.1 lists outputs of FORDISC 3.1 (Jantz and Ousley 2010a). Estimations and tables are presented below specific to the observer's measurements: SLP (Table 4.2), JHE (Table 4.3), ARM (Table 4.4), and the intra-observer/HBD (Table 4.5). Each table also provides the second predicted population group of the unknown individual. Also given are the posterior probabilities for this group and the Mahalanobis' Distances (D^2) in relation to the first predicted group. FORDISC 3.1's sex estimations are in the form of an "M" for "Male" and "F" for "Female" after the population group listed. For example, a White Male and White Female are marked as "WM" and "WF" respectively in FORDISC 3.1.

FORDISC 3.1 regularly provided the same sex estimation for each individual in each of its runs, as well as between skeletal elements. However, for many of these elements, there was very little agreement on what the likely ancestry was for each unknown individual. Between first and second predicted ancestries (see Table 4.1), there were discrepancies in the sex estimated for

some individuals (IND2, IND15, IND16.1 Mandible and Postcrania, IND37 and IND49). In some of these cases, the ancestry predicted was the same with only the sex estimation being different (IND2, IND16.1 Postcrania, IND37, IND49). Predictions made by FORDISC 3.1 from the cranium had greater separation in their Mahalanobis' Distance and their Posterior Probabilities, though this was not the case for the predictions made from mandibular measurements. Secondary predictions for the mandible commonly had posterior probabilities close to the original estimation. These predictions were also close to original predictions in terms of their distances to the centroid.

Table 4.6 gives an overall view of each predicted estimation by all observers. This table also denotes the most frequent estimation for each case. Cases where there was not a conclusive "most frequent ancestry" are marked with "Indeterminate". Individuals who have columns marked with "-" indicate that no data was collected. This occurred with individuals whose crania were present, however, also had taphonomic damage preventing enough measurements (m=4) to be taken in order to run FORDISC 3.1. This also occurred in cases where an observer(s) neglected to take measurements from an unknown individual even when it was possible to do so.

Individuals denoted with the mark ([a]) signify cases where it was noted that the sample sizes of the populations used were not three times the size of the measurements used. As stated in Chapter III, FORDISC 3.1 Help Version 1.48 File deems that the risks of overfitting are not a large concern (Jantz and Ousley 2010b:83). The observer still employed these populations given this information.

Most individuals in Table 4.6 had a consistent agreement between observers in their estimations. Counting skeletal elements (not specific individuals) with estimations predicted by more than one observer (n=29) agreement between observers (i.e. same estimation was predicted

at least half of the time) occurred in 25 cases, with an "Indeterminate" score in only four cases. In 20 of these cases, there was 100% agreement; however, four of these 20 utilized only measurements taken originally by the observer and the secondary measurements taken to assess intra-observer error. Nevertheless, despite disagreements between individual measurements taken by the observers (see Chapter V), the estimation generated did not change in too many instances (n=4) based on cranial measurements (see Table 4.17).

IND1 had the highest level of disagreement between ancestry estimations predicted by the different observers. All four estimations (done by observer, SLP, JHE, and twice by HBD) are different. Section 4.4.1 discusses IND1 further, as this particular individual is noteworthy for observer error as no observer agreed with the initial estimation.

There are cases in the first run of FORDISC 3.1 where the typicalities, particularly the F typicality, were lower than the chosen alpha level of 0.05, and the predicted estimation is very tentative. In one case, IND3, the cranium yielded no result where the F typicality was greater than 0.05. While the individual is listed as "AM[b/c]", this individual can only be documented as male with the ancestry given the cranium remaining unspecified. Other individuals – IND11[b/c] and IND56[b/c] – are also tentative in their ancestry estimations. FORDISC 3.1 does not provide a specific ancestry estimation in any of these cases, instead stating that the unknowns cannot be classified into any of the reference groups used, as they are each too dissimilar to both of them. Such patterns were repeated based on measurements made by other observers, particularly ARM. Any of the ancestries listed below for these tables do not mean to imply that these individuals definitively identified as these estimations in life.

Table 4.1. All Individuals Analyzed using FORDISC 3.1 (Jantz and Ousley 2010a).

Case Number	Element Used	# of Meas. Used	# of Pops. Used	Predicted Ancestry	Posterior Probability	Second Predicted Ancestry	Second Posterior Probability	Distances to Centroid
IND1[a]	Cranium	18	7	HM	0.501	WM	0.268	20.2 --> 21.5
IND2[a]	Cranium	18	13	JF	0.29	JM	0.238	14.8 --> 15.1
IND3	Cranium	-	-	AM[b/c]	-	-	-	-
IND4[a]	Cranium	16	3	AF	0.597	JF[c]	0.328[c]	30.2 --> 31.4
IND5	Cranium	6	12	BF	0.309	WF	0.157	2.0 --> 2.4
IND6	Cranium	13	2	WM	0.857	HM	0.143	14.6 --> 18.2
IND6	Mandible	5	3	GTM	0.453	HM	0.316	7.2 --> 7.9
IND6	Postcranium	11	4	WM	0.969	BM	0.027	4.6 --> 11.8
IND7	Cranium	12	4	BF	0.73	JF	0.206	12.5 --> 15.0
IND8[a]	Cranium	18	12	HF	0.491	BF	0.3	12.7 → 13.7
IND8	Mandible	6	7	WF	0.627	JF	0.217	3.4 --> 5.5
IND9	Cranium	10	3	BF	0.912	JF[c]	0.046[c]	17.1 --> 23.1
IND10	Cranium	17	8	WM	0.572	HM	0.29	20.6 --> 21.9
IND10	Mandible	6	9	GTM	0.221	HM	0.212	5.4 --> 5.5
IND11	Cranium	-	-	AM[b/c]	-	-	-	-
IND11	Mandible	6	9	JM	0.2	GTM	0.199	4.1 --> 4.1
IND12	Cranium	6	12	WM	0.23	CHM	0.207	6.4 --> 6.6
IND13[a]	Cranium	11	13	AM	0.384	BM	0.309	6.5 --> 6.9
IND14	Cranium	17	7	HM	0.269	VM	0.186	21.0 --> 21.8
IND15[a]	Cranium	13	7	AF	0.556	HM	0.128	17.4 --> 20.3
IND16.1	Cranium	-	-	-	-	-	-	-
IND16.1	Mandible	4	9	HM	0.168	BF	0.166	0.3 --> 0.3
IND16.1	Postcranium	11	2	WM	0.92	WF	0.08	13.8 --> 18.7
IND16.2	Postcranium	4	4	BF	0.698	WF	0.252	1.7 --> 3.8
IND23	Cranium	-	-	-	-	-	-	-
IND23	Mandible	4	9	JM	0.275	CHM	0.195	1.7 --> 2.4
IND23	Postcranium	8	3	BM	0.536	WM	0.428	9.4 --> 9.9

Case Number	Element Used	# of Meas. Used	# of Pops. Used	Predicted Ancestry	Posterior Probability	Second Predicted Ancestry	Second Posterior Probability	Distances to Centroid
IND25	Mandible	5	2	JF	0.617	BF	0.383	9.5 --> 10.5
IND26	Mandible	6	9	GTM	0.295	HM	0.205	4.4 --> 5.1
IND27	Mandible	6	9	BM	0.388	GTM	0.178	0.9 --> 2.5
IND28	Cranium	-	-	-	-	-	-	-
IND28	Mandible	-	-	-	-	-	-	-
IND29	Cranium	-	-	-	-	-	-	-
IND29	Mandible	4	9	GTM	0.266	BM	0.199	1.4 --> 2.0
IND30	Mandible	-	-	-	-	-	-	-
IND31	Mandible	-	-	-	-	-	-	-
IND32	Mandible	-	-	-	-	-	-	-
IND33	Mandible	4	3	WM	0.61	JM	0.204	7.3 --> 9.5
IND34	Cranium	-	-	-	-	-	-	-
IND34	Mandible	4	9	JM	0.214	HM	0.164	0.6 --> 1.2
IND35	Cranium	-	-	-	-	-	-	-
IND35	Mandible	4	7	BF	0.395	BM	0.301	2.5 --> 3.0
IND36	Mandible	-	-	-	-	-	-	-
IND37	Cranium	18	13	AM	0.707	AF	0.291	23.3 --> 25.0
IND37	Mandible	-	-	-	-	-	-	-
IND38[a]	Cranium	20	6	HF/HM	0.405/0.402	AF	0.079	27.4/27.4 --> 30.7
IND39	Cranium	9	4	AF	0.929	JF	0.059	11.5 --> 17.0
IND40.1	Mandible	5	9	WF	0.321	JF	0.209	0.8 --> 1.6
IND40.2	Mandible	5	9	CHM	0.371	JM	0.144	0.9 --> 2.8
IND45[a]	Cranium	16	4	BF	0.569	BM	0.295	17.3 --> 18.6
IND46[a]	Cranium	18	12	WM	0.749	AM	0.12	14.4 --> 18.1
IND47[a]	Cranium	13	11	AF	0.472	JF	0.338	6.9 --> 7.5
IND47	Mandible	5	9	WM	0.181	GTM	0.171	1.6 --> 1.7
IND47	Postcranium	8	4	BF	0.775	WF	0.225	9.5 --> 11.9
IND48	Mandible	6	9	BF	0.244	BM	0.232	2.4 --> 2.5
IND49	Mandible	5	9	JM	0.32	JF	0.241	2.5 --> 3.1

53

Case Number	Element Used	# of Meas. Used	# of Pops. Used	Predicted Ancestry	Posterior Probability	Second Predicted Ancestry	Second Posterior Probability	Distances to Centroid
IND50.1[a]	Cranium	11	12	BF	0.592	WF	0.18	10.6 --> 12.9
IND50.2	Mandible	6	9	WF	0.31	BF	0.206	2.9 --> 3.7
IND51[a]	Cranium	11	10	JF	0.475	HF	0.253	9.2 --> 10.5
IND52[a]	Cranium	17	13	AF	0.657	JF	0.108	11.4 --> 15.0
IND53[a]	Cranium	16	5	AF	0.562	AM	0.428	22.7 --> 23.3
IND54	Cranium	16	8	WM	0.991	BM[c]	0.005[c]	26.6 --> 37.0
IND54	Mandible	6	9	JM	0.226	HM	0.16	1.7 --> 2.4
IND55[a]	Cranium	13	11	BF	0.553	WF	0.226	13.3 --> 15.1
IND55	Mandible	5	4	WM	0.576	GTM	0.163	5.0 --> 7.6
IND55	Postcranium	-	-	-	-	-	-	-
IND56	Cranium	16	13	AM[b/c]	0.478[c]	HM[c]	0.263[c]	33.1 --> 34.3
IND56	Mandible	6	9	GTM	0.247	HM	0.225	1.5 --> 1.7
IND57[a]	Cranium	18	6	GTM	0.591	HF	0.211	20.7 --> 22.8

[a]Not all sample sizes of reference populations are three times the size of the number of measurements used.
[b]Context and patterns observed through FORDISC 3.1 runs suggests American Indian ancestry.
[c]Typicalities < 0.05

Table 4.2. Individuals Measured by SLP for Inter-observer Error using FORDISC 3.1 (Jantz and Ousley 2010a).

Case Number	Element Used	# of Meas. Used	# of Pops. Used	Predicted Ancestry	Posterior Probability	Second Predicted Ancestry	Second Posterior Probability	Distances to Centroid
IND1[a]	Cranium	18	12	GTM	0.364	HM	0.2	14.4 --> 15.6
IND6[a]	Cranium	18	13	WM	0.988	HM[b]	0.007[b]	28.6 --> 38.5
IND6	Mandible	5	3	GTM	0.417	HM	0.317	4.8 --> 5.3
IND12	Cranium	6	13	WM	0.163	JM	0.137	1.5 --> 1.8
IND14[a]	Cranium	18	11	VM	0.52	BF	0.111	20.5 --> 23.6
IND15[a]	Cranium	13	11	AF	0.526	HM	0.118	13.4 --> 16.4
IND27	Mandible	6	9	BM	0.47	BF	0.169	2.3 --> 4.3
IND33	Mandible	4	6	WM	0.305	JM	0.226	4.4 --> 5.0
IND40.1	Mandible	5	9	WF	0.281	WM	0.19	1.2 --> 2.0
IND40.2	Mandible	5	9	CHM	0.341	HM	0.158	0.9 --> 2.4
IND51	Cranium	11	6	JF	0.576	HF	0.138	13.6 --> 16.4

[a]Not all sample sizes of reference populations are three times the size of the number of measurements used.
[b]Typicalities < 0.05.

Table 4.3. Individuals Measured by JHE for Inter-observer Error using FORDISC 3.1 (Jantz and Ousley 2010a).

Case Number	Element Used	# of Meas. Used	# of Pops. Used	Predicted Ancestry	Posterior Probability	Second Predicted Ancestry	Second Posterior Probability	Distances to Centroid
IND1[a]	Cranium	18	11	BM	0.487	CHM	0.147	14.5 --> 16.9
IND6[a]	Cranium	18	13	WM	0.962	HM[b]	0.019[b]	28.2 --> 36.0
IND12	Cranium	6	13	CHM	0.22	WM	0.158	5.6 --> 6.2
IND14[a]	Cranium	17	8	WM	0.296	WF	0.229	16.7 --> 17.2
IND15[a]	Cranium	13	11	AF	0.273	HM	0.163	17.6 --> 18.6
IND27	Mandible	6	9	BM	0.504	GTM	0.131	2.3 --> 5.0
IND33	Mandible	4	5	WM	0.313	JM	0.218	5.0 --> 5.9
IND40.1	Mandible	5	9	WF	0.258	JF	0.183	0.4 --> 1.1
IND40.2	Mandible	5	9	CHM	0.363	HM	0.135	0.9 --> 2.9
IND51[a]	Cranium	11	10	JF	0.816	HF	0.121	8.2 --> 12.1

[a]Not all sample sizes of reference populations are three times the size of the number of measurements used.
[b]Typicalities < 0.05

Table 4.4. Individuals Measured by ARM for Inter-observer Error using FORDISC 3.1 (Jantz and Ousley 2010a).

Case Number	Element Used	# of Meas. Used	# of Pops. Used	Predicted Ancestry	Posterior Probability	Second Predicted Ancestry	Second Posterior Probability	Distances to Centroid
IND2[a]	Cranium	19	11	JF	0.301	GTM	0.231	17.7 --> 18.3
IND3[c]	Cranium	-	-	AM[b/c]	-	-	-	-
IND4[a]	Cranium	16	3	AF	0.597	JF[c]	0.328[c]	30.2 --> 31.4
IND11	Cranium	-	-	AM[b/c]	-	-	-	-
IND13[a]	Cranium	11	13	AM	0.384	BM	0.309	6.5 --> 6.9
IND37	Cranium	18	13	AM	0.707	AF	0.291	23.3 --> 25.0
IND39	Cranium	9	4	AF	0.929	JF	0.059	11.5 --> 17.0
IND46[a]	Cranium	18	13	WM	0.776	AM	0.094	13.1 --> 17.3
IND47[a]	Cranium	13	11	AF	0.472	JF	0.338	6.9 --> 7.5
IND52[a]	Cranium	17	13	AF	0.583	JF	0.137	10.3 --> 13.2
IND53[a]	Cranium	16	11	AF	0.949	AM	0.051	25.8 --> 31.6
IND57[a]	Cranium	18	6	GTM	0.591	HF	0.211	20.7 --> 22.8

[a]Not all sample sizes of reference populations are three times the size of the number of measurements used.
[b]Context and patterns observed through FORDISC 3.1 runs suggests American Indian ancestry.
[c]Typicalities < 0.05.

Table 4.5. Individuals Measured by HBD for Intra-observer Error using FORDISC 3.1 (Jantz and Ousley 2010a).

Case Number	Element Used	# of Meas. Used	# of Pops. Used	Predicted Ancestry	Posterior Probability	Second Predicted Ancestry	Second Posterior Probability	Distances to Centroid
IND1[a]	Cranium	17	13	WM	0.328	HM	0.282	6.4 --> 7.1
IND2[a]	Cranium	18	13	JF	0.367	JM	0.158	15.5 --> 17.2
IND3[a]	Cranium	20	6	AM	0.802	GTM	0.173	21.1 --> 24.2
IND4[a]	Cranium	18	4	AF	0.6	AM	0.186	28.5 --> 30.9
IND6	Cranium	19	3	WM	0.711	HM	0.26	23.9 --> 25.9
IND6	Mandible	5	3	GTM	0.454	HM	0.337	6.0 --> 6.6
IND11[a]	Cranium	-	-	AM[b/c]	-	-	-	-
IND11	Mandible	6	9	CHM	0.276	JM	0.232	2.8 --> 3.2
IND12	Cranium	6	13	WM	0.166	BM	0.16	1.2 --> 1.3
IND13[a]	Cranium	11	13	BM	0.363	AM	0.278	8.6 --> 9.1
IND14[a]	Cranium	18	11	VM	0.473	HM	0.146	20.5 --> 22.9
IND15[a]	Cranium	14	5	AF	0.645	CHM[c]	0.129[c]	24.4 --> 27.7
IND27	Mandible	6	9	BM	0.492	GTM	0.152	1.3 --> 3.7
IND33	Mandible	4	6	WM	0.305	JM	0.226	4.4 --> 5.0
IND38[a]	Cranium	17	13	HM	0.433	CHM	0.076	16.3 --> 19.8
IND40.1	Mandible	5	9	WF	0.213	WM	0.178	2.6 --> 3.0
IND40.2	Mandible	5	9	CHM	0.269	WM	0.179	1.7 --> 2.5
IND46[a]	Cranium	18	13	WM	0.778	AM	0.057	13.0 --> 18.3
IND47[a]	Cranium	12	9	JF	0.648	BF	0.225	8.2 --> 10.3
IND47	Mandible	5	9	HM	0.163	WM	0.131	1.1 --> 1.2
IND47	Postcranium	8	4	BF	0.752	WF	0.248	11.5 --> 13.7
IND50.1[a]	Cranium	11	13	BF	0.69	WF	0.09	9.3 --> 13.4
IND50.2	Mandible	6	9	BF	0.255	WF	0.227	2.8 --> 3.1
IND51	Cranium	11	6	JF	0.796	HF	0.085	11.3 --> 15.7
IND52[a]	Cranium	17	13	AF	0.437	HM	0.176	13.1 --> 14.9
IND53[a]	Cranium	18	4	AF	0.637	AM	0.356	15.5 --> 16.7
IND57[a]	Cranium	18	7	GTM	0.433	HF	0.265	21.6 --> 22.5

[a]Not all sample sizes of reference populations are three times the size of the number of measurements used.
[b]Context and patterns observed through FORDISC 3.1 runs suggests American Indian ancestry.
[c]Typicalities < 0.05.

Table 4.6. Frequency of First Predicted Ancestry Estimations using each Observer's Measurements using FORDISC 3.1 (Jantz and Ousley 2010a).

Case Number	Element Used	Original Ancestry	SLP Ancestry	JHE Ancestry	ARM Ancestry	INTRA Ancestry	Most Common Ancestry
IND1[a]	Cranium	HM	GTM	BM	-	WM	Indeterminate
IND2[a]	Cranium	JF	-	-	JF	JF	JF (3/3)
IND3	Cranium	AM[b/c]	-	-	AM[b/c]	AM	AM[b] (3/3)
IND4[a]	Cranium	AF	-	-	AF	AF	AF (3/3)
IND5	Cranium	BF	-	-	-	-	BF (1/1)
IND6	Cranium	WM	WM	WM	-	WM	WM (4/4)
IND6	Mandible	GTM	GTM	-	-	GTM	GTM (3/3)
IND6	Postcranium	WM	-	-	-	-	WM (1/1)
IND7	Cranium	BF	-	-	-	-	BF (1/1)
IND8[a]	Cranium	HF	-	-	-	-	HF (1/1)
IND8	Mandible	WF	-	-	-	-	WF (1/1)
IND9	Cranium	BF	-	-	-	-	BF (1/1)
IND10	Cranium	WM	-	-	-	-	WM (1/1)
IND10	Mandible	GTM	-	-	-	-	GTM (1/1)
IND11	Cranium	AM[b/c]	-	-	AM[b/c]	VM	AM[b] (2/3)
IND11	Mandible	JM	-	-	-	CHM	Indeterminate
IND12	Cranium	WM	WM	CHM	-	WM	WM (3/4)
IND13[a]	Cranium	AM	-	-	AM	BM	AM (2/3)
IND14	Cranium	HM	VM	WM	-	VM	VM (2/4)
IND15[a]	Cranium	AF	AF	AF	-	AF	AF (4/4)
IND16.1	Cranium	-	-	-	-	-	-
IND16.1	Mandible	HM	-	-	-	-	HM (1/1)
IND16.1	Postcranium	WM	-	-	-	-	WM (1/1)
IND16.2	Postcranium	BF	-	-	-	-	BF (1/1)
IND23	Cranium	-	-	-	-	-	-
IND23	Mandible	JM	-	-	-	-	JM (1/1)
IND23	Postcranium	BM	-	-	-	-	BM (1/1)

Case Number	Element Used	Original Ancestry	SLP Ancestry	JHE Ancestry	ARM Ancestry	INTRA Ancestry	Most Common Ancestry
IND25	Mandible	JF	-	-	-	-	JF (1/1)
IND26	Mandible	GTM	-	-	-	-	GTM (1/1)
IND27	Mandible	BM	BM	BM	-	BM	BM (4/4)
IND28	Cranium	-	-	-	-	-	-
IND28	Mandible	-	-	-	-	-	-
IND29	Cranium	-	-	-	-	-	-
IND29	Mandible	GTM	-	-	-	-	GTM (1/1)
IND30	Mandible	-	-	-	-	-	-
IND31	Mandible	-	-	-	-	-	-
IND32	Mandible	-	-	-	-	-	-
IND33	Mandible	WM	WM	WM	-	WM	WM (4/4)
IND34	Cranium	-	-	-	-	-	-
IND34	Mandible	JM	-	-	-	-	JM (1/1)
IND35	Cranium	-	-	-	-	-	-
IND35	Mandible	BF	-	-	-	-	BF (1/1)
IND36	Mandible	-	-	-	-	-	-
IND37	Cranium	AM[b/c]	-	-	AM[b/c]	-	AM[b] (2/2)
IND37	Mandible	-	-	-	-	-	-
IND38[a]	Cranium	HF/HM	-	-	-	HM	HM (2/2)
IND39	Cranium	AF	-	-	AF	-	AF (2/2)
IND40.1	Mandible	WF	WF	WF	-	WF	WF (4/4)
IND40.2	Mandible	CHM	CHM	CHM	-	CHM	CHM (4/4)
IND45[a]	Cranium	BF	-	-	-	-	BF (1/1)
IND46[a]	Cranium	WM	-	-	WM	WM	WM (3/3)
IND47[a]	Cranium	AF	-	-	AF	JF	AF (2/3)
IND47	Mandible	WM	-	-	-	HM	Indeterminate
IND47	Postcranium	BF	-	-	-	BF	BF (2/2)
IND48	Mandible	BF	-	-	-	-	BF (1/1)
IND49	Mandible	JM	-	-	-	-	JM (1/1)
IND50.1[a]	Cranium	BF	-	-	-	BF	BF (2/2)

Case Number	Element Used	Original Ancestry	SLP Ancestry	JHE Ancestry	ARM Ancestry	INTRA Ancestry	Most Common Ancestry
IND50.2	Mandible	WF	-	-	-	BF	Indeterminate
IND51[a]	Cranium	JF	JF	JF	-	JF	JF (4/4)
IND52[a]	Cranium	AF	-	-	AF	AF	AF (3/3)
IND53[a]	Cranium	AF	-	-	AF	AF	AF (3/3)
IND54	Cranium	WM	-	-	-	-	WM (1/1)
IND54	Mandible	JM	-	-	-	-	JM (1/1)
IND55[a]	Cranium	BF	-	-	-	-	BF (1/1)
IND55	Mandible	WM	-	-	-	-	WM (1/1)
IND55	Postcranium	-	-	-	-	-	-
IND56	Cranium	AM[b/c]	-	-	-	-	AM[b] (1/1)
IND56	Mandible	GTM	-	-	-	-	GTM (1/1)
IND57[a]	Cranium	GTM	-	-	GTM	GTM	GTM (3/3)

[a]Not all sample sizes of reference populations are three times the size of the number of measurements used.
[b]Context and patterns observed through FORDISC 3.1 runs suggests American Indian ancestry.
[c]Typicalities < 0.05.

4.2.2 Ancestry Estimation from the Mandible

Using Berg (2015), ancestry estimations based off composite world populations using the mandible were generated. Table 4.7 provides outputs from the online GUI, (hu)MANid using linear discriminant analyses (LDA) based upon the metric data (MET) obtained through measuring. Table 4.8 uses these same measurements but instead employs a mixture discriminant analysis (MDA). In all forms of observation (metric, morphoscopic and morphometroscopic), MDA results often had a higher cross-validation than LDA even though they regularly relied on the same population groups as LDA. The observer noticed this trend while running all data through the GUI. Although more populations are available as results than FORDISC 3.1, some of them (e.g. Southeast/Northeast Asian) do not seem likely given the overall context of the sample.

The second-most likely estimation and its posterior probability are also included in Tables 4.7 and 4.8 along with the Euclidean Distance in respect to the predicted ancestry group. In only several cases of the 28 mandibles did the posterior probability of the predicted group and second group correspond with the Euclidean Distance value (LDA$_{MET}$ n=11; MDA$_{MET}$ n=6), and there were frequently other population groups that were listed as being "closer" in distance to the unknown individual (LDA$_{MET}$ n=9; MDA$_{MET}$ n=16). This was seen regardless of the type of variable (e.g. measurement, morphoscopic trait or morphometroscopic feature) being used.

Tables 4.9 and 4.10 provide ancestry estimations when only morphoscopic data (MORPH) are used. The linear discriminant analysis outputs are in Table 4.9 and mixture discriminant analysis outputs are in Table 4.10. There were only two cases (IND28 and IND30) where an estimation by the GUI was impossible due taphonomic damage affecting observation of the morphoscopic features.

There were again discrepancies with the "Distances to Centroid" portion where the order

of likelihood represented in the "Posterior Probabilities" columns correspond infrequently with

the distance (LDA$_{MORPH}$ n= 5; MDA$_{MORPH}$ n=9). There were also many instances where the

Euclidean Distance predicted another ancestry to be more likely (LDA$_{MORPH}$ n=15; MDA$_{MORPH}$

n=14).

Similarly, the morphometroscopic data (MM) – the combination of the metric data and

morphoscopic data – were analyzed through an LDA (Table 4.11) and an MDA (Table 4.12).

Occurrences of correspondence between the predicted "Posterior Probabilities" and the

"Distances to Centroid" were few with these data as well (LDA$_{MM}$ n=12; MDA$_{MM}$ n=2). Using

LDA with these data had the highest rate of correspondence while using MDA had the lowest,

regardless of data type (metrics, morphoscopics and morphometroscopics). Again, there were

several cases where the Euclidean Distance indicates that perhaps a different group is more likely

than those predicted with the "Posterior Probabilities" (LDA$_{MM}$ n=12; MDA$_{MM}$ n=20). Using the

morphometroscopic variable resulted in the highest rate of occurrence where MDA is used and

the distance indicates a different ancestry between all data types.

Table 4.13 lists the frequency of estimations using all above analyses. This also includes

estimations using measurements taken by the inter-observers (SLP and JHE) and the intra-

observer (HBD). Tables specific to inter-/intra-observer LDAs' and MDAs' outputs can be found

in the Appendix B with the same type of information provided in Table 4.7-Table 4.13.

Agreement where at least half of the ancestry estimations were the same between users

occurred in 25 instances, with only three individuals (IND11, IND16.1 and IND56) scored as

"Indeterminate". All of these cases involved measurements only taken by the observer (HBD),

with IND11 measured again by observer for intra-observer error. IND11 and IND56 mark

important cases where observer error for specific measurements may have impacted the

estimation. Chapter V discusses these two individuals at greater length. The inconsistencies with

IND16.1 are harder to explain. All but two of the runs done with this individual (Tables 4.7-4.13)

– despite data type – had different ancestries predicted through the Euclidean Distance.

All tables that have individuals marked with "-" signify cases where an analysis could not

be conducted based on the degree of taphonomic damage/preservation obscuring necessary

landmarks. Ancestry estimations for each table and the most common ancestry listed in Table

4.13 do not necessarily correctly estimate how any of the individuals identified in life.

Table 4.7. (hu)MANid (Berg and Kenyhercz 2017) Outputs using Measurements and Linear Discriminant Analysis (LDA).

Case Number	# of Meas. Used	# of Pops. Used	Predicted Ancestry	Posterior Probability	Second Predicted Ancestry	Posterior Probability	Distances to Centroid
IND6	9	3	WM	0.897	NEAM	0.086	2.105 --> 3.022
IND8	8	6	WF	0.642	WM	0.268	0.826 --> 2.271
IND10	8	7	WM	0.535	HISPM	0.237	1.608 --> 2.225
IND11	9	9	HISPM	0.25	AIM	0.224	AIM(0.386)--->0.960
IND16.1	6	10	HISPM	0.327	NEAM	0.125	0.237 --> 0.275
IND23	5	11	AIM	0.217	HISPM	0.163	BM(0.204) --> 0.221
IND25	8	3	AIF	0.563	BF	0.266	0.680 --> 1.401
IND26	8	8	BM	0.295	HISPM	0.252	0.322 --> AIM(0.544)
IND27	9	8	HISPM	0.427	BM	0.173	0.755 --> 0.789
IND28	5	9	BF	0.558	AIF	0.346	1.544 --> 1.646
IND29	6	4	AIF	0.494	HISPM	0.234	0.541 --> AIM(1.421)
IND30	4	7	BM	0.325	HISPM	0.282	AIM(0.232) --> 0.240
IND31	4	5	AIF	0.733	BF	0.1	2.015 --> 2.515
IND32	4	3	BF	0.625	AIF	0.291	1.084 --> 1.645
IND33	5	8	HISPM	0.259	NEAM	0.251	WM(0.284)--->0.662
IND34	6	9	AIM	0.399	BM	0.205	BM(0.282) --> 0.364
IND35	7	8	BF	0.586	AIF	0.358	AIF(1.918) --> 2.044
IND36	4	12	HISPF	0.498	SEAF	0.175	1.062 --> WF(1.229)
IND37	3	10	SEAM	0.213	AIM	0.133	NEAF(0.435)--->0.876
IND40.1	8	9	BF	0.314	SEAF	0.296	0.418 --> AIF(1.184)
IND40.2	8	4	AIM	0.634	HISPM	0.249	1.294 --> NEAM(1.705)
IND47	6	10	AIF	0.618	BF	0.138	0.523 --> 1.049
IND48	9	5	BF	0.389	AIF	0.244	AIF(0.423) --> 0.638
IND49	8	3	SEAF	0.58	SEAM	0.274	0.330 --> 1.267
IND50.2	9	6	WM	0.397	HISPM	0.316	0.267 --> SEAM(1.365)
IND54	7	9	WM	0.242	HISPM	0.177	0.431 --> SEAF(0.614)
IND55	8	6	HISPM	0.394	WM	0.215	WM(0.313)--->1.194
IND56	9	4	AIM	0.417	HISPM	0.256	0.729 --> BM(0.812)

Table 4.8. (hu)MANid Outputs (Berg and Kenyhercz 2017) using Measurements and Mixture Discriminant Analysis (MDA).

Case Number	# of Meas. Used	# of Pops. Used	Predicted Ancestry	Posterior Probability	Second Predicted Ancestry	Posterior Probability	Distances to Centroid
IND6	9	3	WM	0.858	SEAM	0.105	2.646 --> 3.421
IND8	8	6	WM	0.778	WF	0.191	WF(1.513)--->3.147
IND10	8	7	WM	0.79	HISPM	0.063	HISPM(1.390)--->1.605
IND11	9	9	HISPM	0.326	AIM	0.202	AIM(0.350)--->1.106
IND16.1	6	10	HISPM	0.273	WM	0.194	NEAM(0.377) --> 0.507
IND23	5	11	WM	0.336	SEAM	0.225	HISPM(1.108)--->1.444
IND25	8	3	BF	0.476	AIF	0.287	0.433 --> 1.694
IND26	8	8	WM	0.429	BM	0.261	SEAF(1.006)--->1.956
IND27	9	8	HISPM	0.333	BM	0.201	AIF(0.557)--->1.276
IND28	5	9	BF	0.847	AIF	0.107	SEAF(3.017) --> 3.439
IND29	6	4	BM	0.456	AIM	0.216	AIM(2.218)--->3.240
IND30	4	7	BM	0.424	HISPM	0.298	0.895 --> AIM(1.064)
IND31	4	5	AIF	0.787	BF	0.145	2.748 --> BM(3.041)
IND32	4	3	BF	0.597	AIF	0.263	AIF(1.440) --> 1.517
IND33	5	8	HISPM	0.445	NEAM	0.188	NEAM(1.121)--->2.195
IND34	6	9	WM	0.25	BM	0.233	HISPM(0.274)--->1.054
IND35	7	8	BF	0.763	AIF	0.184	1.445 --> 2.049
IND36	4	12	HISPF	0.822	SEAM	0.049	WF(0.334)--->1.316
IND37	3	10	SEAM	0.544	NEAM	0.201	3.249 --> AIM(3.301)
IND40.1	8	9	BF	0.396	SEAF	0.288	0.930 --> AIF(1.475)
IND40.2	8	4	AIM	0.597	NEAM	0.268	1.031 --> 1.464
IND47	6	10	AIF	0.439	BF	0.173	0.547 --> 0.757
IND48	9	5	BF	0.423	HISPM	0.207	0.298 --> AIF(0.780)
IND49	8	3	SEAM	0.509	SEAF	0.434	SEAF(1.135) --> 1.710
IND50.2	9	6	WM	0.841	HISPM	0.087	2.120 --> SEAF(2.367)
IND54	7	9	WM	0.473	SEAM	0.176	0.185 --> 0.765
IND55	8	6	WM	0.552	NEAM	0.173	NEAM(0.815)--->2.326
IND56	9	4	HISPM	0.437	BM	0.33	AIM(0.851)--->1.958

Table 4.9. (hu)MANid (Berg and Kenyhercz 2017) Outputs using Morphoscopics and Linear Discriminant Analysis (LDA).

Case Number	# of Traits Used	# of Pops. Used	Predicted Ancestry	Posterior Probability	Second Predicted Ancestry	Posterior Probability	Distances to Centroid
IND6	6	6	WM	0.851	AIM	0.051	0.851 --> BM(3.511)
IND8	6	7	BF	0.501	BM	0.264	2.559 --> 2.758
IND10	6	9	WM	0.532	WF	0.275	BM(1.302) --> 1.548
IND11	6	3	BF	0.614	WF	0.292	0.830 --> 1.475
IND16.1	6	8	BM	0.371	WF	0.194	WF(0.560) --> 0.586
IND23	6	8	AIF	0.284	HISPM	0.252	HISPM(1.442)--->1.942
IND25	6	8	HISPM	0.313	BM	0.175	0.993 --> WM(1.155)
IND26	6	7	BM	0.371	BF	0.216	BF(0.803) --> 1.105
IND27	6	8	HISPM	0.295	BM	0.172	WM(1.148)--->1.470
IND28	-	-	-	-	-	-	-
IND29	5	3	WM	0.498	WF	0.346	0.719 --> BM(1.150)
IND30	-	-	-	-	-	-	-
IND31	5	10	NEAF	0.216	SEAM	0.156	SEAM(0.504) --> 0.604
IND32	5	4	BF	0.365	SEAF	0.289	0.442 --> 1.303
IND33	5	8	AIM	0.353	AIF	0.143	HISPM(0.679) --> 0.735
IND34	5	8	BF	0.651	BM	0.149	2.939 --> BM(3.420)
IND35	5	4	BM	0.35	BF	0.329	1.044 --> 1.082
IND36	5	9	WM	0.605	AIM	0.088	1.320 --> WF(1.877)
IND37	6	8	WM	0.265	AIM	0.247	1.352 --> BM(1.673)
IND40.1	6	9	AIF	0.19	SEAM	0.18	1.069 --> 1.293
IND40.2	6	7	BM	0.412	BF	0.271	BF(2.175) --> 2.304
IND47	6	9	AIF	0.221	SEAF	0.147	SEAM(0.124)--->0.789
IND48	6	9	WM	0.518	BM	0.188	BM(2.553) --> 2.639
IND49	6	9	SEAM	0.201	AIM	0.188	AIF(0.959)--->1.288
IND50.2	6	9	WF	0.319	WM	0.251	WM(0.959) --> 1.466
IND54	6	8	AIF	0.354	HISPM	0.152	HISPM(0.266)--->0.596
IND55	6	10	WM	0.686	BM	0.111	BM(1.907) --> 2.057
IND56	6	10	WM	0.548	WF	0.194	BM(0.746) --> 0.985

Table 4.10. (hu)MANid (Berg and Kenyhercz 2017) Outputs using Morphoscopics and Mixture Discriminant Analysis (MDA).

Case Number	# of Traits Used	# of Pops. Used	Predicted Ancestry	Posterior Probability	Second Predicted Ancestry	Posterior Probability	Distances to Centroid
IND6	6	3	WM	0.959	WF	0.041	3.648 --> 5.126
IND8	6	7	BF	0.812	BM	0.09	1.161 --> 1.912
IND10	6	9	WM	0.784	WF	0.199	3.153 --> 4.001
IND11	6	3	BF	0.55	WF	0.345	1.295 --> BM(2.047)
IND16.1	6	8	WM	0.409	WF	0.315	3.463 --> 4.154
IND23	6	8	HISPM	0.314	WM	0.231	AIF(0.355)--->0.992
IND25	6	8	NEAM	0.452	AIF	0.168	SEAM(0.356)--->0.578
IND26	6	7	WF	0.455	BM	0.259	BF(0.502)--->2.315
IND27	6	8	HISPM	0.338	WM	0.3	WM(0.722)--->2.615
IND28	-	-	-	-	-	-	-
IND29	5	3	WM	0.736	BM	0.14	3.362 --> WF(4.461)
IND30	-	-	-	-	-	-	-
IND31	5	10	SEAM	0.501	WM	0.175	WM(0.749)--->1.924
IND32	5	4	SEAF	0.707	BF	0.132	WF(2.876)--->3.429
IND33	5	8	WM	0.262	AIM	0.239	AIF(0.778)--->2.846
IND34	5	8	BF	0.652	BM	0.179	3.603 --> 4.360
IND35	5	4	BF	0.412	SEAM	0.332	WF(1.751)--->2.134
IND36	5	9	WM	0.626	SEAM	0.133	1.341 --> WF(2.641)
IND37	6	8	WM	0.698	BM	0.09	AIF(0.843)--->3.130
IND40.1	6	9	SEAM	0.464	SEAF	0.109	AIF(1.052)--->2.212
IND40.2	6	7	WF	0.518	BM	0.247	BF(0.485)--->2.150
IND47	6	9	SEAM	0.221	WF	0.153	SEAF(1.031)--->1.554
IND48	6	9	WM	0.888	WF	0.067	1.939 --> 2.514
IND49	6	9	SEAM	0.474	WM	0.141	AIF(0.869)--->1.972
IND50.2	6	9	WM	0.729	WF	0.126	2.124 --> 3.219
IND54	6	8	BM	0.173	HISPM	0.17	AIF(0.543)--->1.501
IND55	6	10	WM	0.944	WF	0.039	3.458 --> 4.247
IND56	6	10	WM	0.824	WF	0.095	1.434 --> 2.621

Table 4.11. (hu)MANid (Berg and Kenyhercz 2017) Outputs using Morphometroscopics and Linear Discriminant Analysis (LDA).

Case Number	# of Feat. Used	# of Pops. Used	Predicted Ancestry	Posterior Probability	Second Predicted Ancestry	Posterior Probability	Distances to Centroid
IND6	15	5	WM	0.931	HISPM	0.039	2.512 --> BM(3.007)
IND8	14	3	WF	0.584	BF	0.41	2.338 --> 2.484
IND10	14	8	WM	0.914	BM	0.056	SEAM(0.099) --> 0.211
IND11	15	7	BM	0.293	BF	0.259	0.778 --> 1.567
IND16.1	12	10	BM	0.293	WM	0.233	SEAF(0.578)--->1.676
IND23	11	5	HISPM	0.364	AIM	0.285	AIM(0.412)--->0.887
IND25	14	3	AIF	0.887	SEAF	0.068	0.583 --> 2.341
IND26	14	3	BM	0.715	AIF	0.236	1.184 --> 1.902
IND27	15	7	HISPM	0.583	BM	0.231	BM(0.861) --> 1.252
IND28	8	5	BF	0.702	AIF	0.244	1.571 --> 2.086
IND29	11	6	HISPM	0.277	BM	0.211	BM(0.419)--->0.850
IND30	7	8	BM	0.319	BF	0.244	AIF(0.504)--->0.942
IND31	9	6	AIF	0.742	BM	0.113	BM(0.884) --> 0.994
IND32	9	5	BF	0.975	BM	0.01	2.231 --> 3.542
IND33	10	7	HISPM	0.327	NEAM	0.197	WM(0.359) --> 0.491
IND34	11	9	BM	0.418	BF	0.36	BF(0.779)--->1.492
IND35	12	6	BF	0.931	AIF	0.038	1.670 --> 2.280
IND36	9	8	WM	0.629	SEAF	0.118	WF(0.042)--->1.360
IND37	9	7	AIM	0.419	BM	0.126	0.135 --> 0.251
IND40.1	14	9	SEAF	0.494	AIF	0.222	0.175 --> BM(1.322)
IND40.2	14	4	AIM	0.55	HISPM	0.219	0.253 --> NEAM(1.494)
IND47	12	9	AIF	0.713	BF	0.09	0.795 --> 1.253
IND48	15	7	BF	0.509	BM	0.321	0.915 --> SEAF(1.032)
IND49	14	3	SEAF	0.662	SEAM	0.297	0.728 --> 1.462
IND50.2	15	5	WM	0.786	HISPM	0.131	0.693 --> 2.067
IND54	13	8	HISPM	0.288	WM	0.183	NEAM(0.194) --> 0.461
IND55	14	5	WM	0.884	HISPM	0.072	1.227 --> 2.527
IND56	15	7	AIM	0.4	BM	0.248	BM(0.677)--->1.673

Table 4.12. (hu)MANid (Berg and Kenyhercz 2017) Outputs using Morphometroscopics and Mixture Discriminant Analysis (MDA).

Case Number	# of Feat. Used	# of Pops. Used	Predicted Ancestry	Posterior Probability	Second Predicted Ancestry	Posterior Probability	Distances to Centroid
IND6	15	5	WM	0.984	HISPM	0.008	2.904 --> BM(4.068)
IND8	14	3	BF	0.524	WF	0.476	2.518 --> 2.639
IND10	14	8	WM	0.986	BM	0.011	2.004 --> HISPM(2.883)
IND11	15	7	BM	0.346	AIM	0.223	HISPM(0.354) --> 0.565
IND16.1	12	10	WM	0.361	BM	0.251	SEAM(0.175)--->0.729
IND23	11	5	HISPM	0.376	NEAM	0.2	BM(1.194)--->1.560
IND25	14	3	AIF	0.679	SEAF	0.297	0.532 --> BF(2.502)
IND26	14	4	BM	0.741	AIF	0.155	AIF(0.794) --> 1.749
IND27	15	7	HISPM	0.618	BM	0.224	AIF(0.586) --> 1.079
IND28	8	5	BF	0.692	AIF	0.253	SEAF(2.845)--->3.308
IND29	11	6	WM	0.299	HISPM	0.209	HISPM(0.651)--->2.483
IND30	7	8	WM	0.786	BM	0.124	NEAM(1.510)--->2.141
IND31	9	6	AIF	0.763	BF	0.133	BM(1.150) --> 1.494
IND32	9	5	BF	0.985	WF	0.011	2.507 --> AIF(3.159)
IND33	10	7	HISPM	0.594	WM	0.173	AIM(1.470)--->2.564
IND34	11	9	BF	0.531	WM	0.298	NEAM(0.811)--->1.094
IND35	12	6	BF	0.97	BM	0.013	1.472 --> AIF(2.297)
IND36	9	8	WM	0.83	WF	0.069	WF(0.661)--->1.611
IND37	9	7	SEAM	0.367	NEAM	0.274	AIM(0.906)--->1.509
IND40.1	14	9	SEAF	0.46	WM	0.205	AIF(0.654)--->1.031
IND40.2	14	4	BM	0.741	NEAM	0.112	AIM(0.697)--->2.487
IND47	12	9	AIF	0.608	BF	0.165	SEAF(0.231) --> 0.798
IND48	15	7	BF	0.643	WM	0.199	WF(1.466) --> 1.581
IND49	14	3	SEAF	0.516	SEAM	0.43	0.925 --> 2.025
IND50.2	15	5	WM	0.978	HISPM	0.012	HISPM(2.215)--->2.358
IND54	13	8	WM	0.394	HISPM	0.224	HISPM(0.412)--->0.572
IND55	14	5	WM	0.927	HISPM	0.028	1.257 --> BM(2.016)
IND56	15	7	SEAM	0.552	WM	0.219	WM(1.147) --> 2.072

Table 4.13. Frequency of Most Predicted Ancestry with All Analyses and Observers using (hu)MANid (Berg and Kenyhercz 2017).

Case Number	LDA MET.	MDA MET.	LDA MORPH.	MDA MORPH.	LDA MM.	MDA MM.	LDA SLP	MDA SLP	LDA JHE	MDA JHE	LDA INTRA	MDA INTRA	Most Common Ancestry
IND6	WM	WM	WM	WM	WM	WM	WM	WM	-	-	AIM	WM	WM (9/10)
IND8	WF	WM	BF	BF	WF	BF	-	-	-	-	-	-	BF (3/5)
IND10	WM	WM	WM	WM	WM	WM	-	-	-	-	-	-	WM (6/6)
IND11	HISPM	HISPM	BF	BF	BM	BM	-	-	-	-	AIM	NEAM	Indeterminate
IND16.1	HISPM	HISPM	BM	WM	BM	WM	-	-	-	-	-	-	Indeterminate
IND23	AIM	WM	AIF	HISPM	HISPM	HISPM	-	-	-	-	-	-	HISPM (3/6)
IND25	AIF	BF	HISPM	NEAM	AIF	AIF	-	-	-	-	-	-	AIF (3/6)
IND26	BM	WM	BM	WF	BM	BM	-	-	-	-	-	-	BM (4/6)
IND27	HISPM	HISPM	HISPM	HISPM	HISPM	HISPM	HISPM	HISPM	HISPM	HISPM	HISPM	HISPM	HISPM (12/12)
IND28	BF	BF	N/A	N/A	BF	BF	-	-	-	-	-	-	BF (4/4)
IND29	AIF	BM	WM	WM	HISPM	WM	-	-	-	-	-	-	WM (3/6)
IND30	BM	BM	N/A	N/A	BM	WM	-	-	-	-	-	-	BM (3/4)
IND31	AIF	AIF	NEAF	SEAM	AIF	AIF	-	-	-	-	-	-	AIF (4/6)
IND32	BF	BF	BF	SEAF	BF	BF	HISPF	BF	HISPF	BF	-	-	BF (7/10)
IND33	HISPM	HISPM	AIM	WM	HISPM	HISPM	NEAM	HISPM	NEAM	HISPM	AIM	BM	HISPM (6/12)
IND34	AIM	WM	BF	BF	BM	BF	-	-	-	-	-	-	BF (3/6)
IND35	BF	BF	BM	BF	BF	BF	-	-	-	-	-	-	BF (5/6)
IND36	HISPF	HISPF	WM	WM	WM	WM	-	-	-	-	-	-	WM (4/6)
IND37	SEAM	SEAM	WM	WM	AIM	SEAM	-	-	-	-	-	-	SEAM (3/6)
IND40.1	BF	BF	AIF	SEAM	SEAF	SEAF	AIF	SEAF	AIF	AIF	AIF	AIF	AIF (6/12)
IND40.2	AIM	AIM	BM	WF	AIM	BM	AIM	AIM	AIM	AIM	AIM	AIM	AIM (9/12)
IND47	AIF	AIF	AIF	SEAM	AIF	AIF	-	-	-	-	AIF	AIF	AIF (7/8)
IND48	BF	BF	WM	WM	BF	BF	-	-	-	-	-	-	BF (4/6)
IND49	SEAF	SEAM	SEAM	SEAM	SEAF	SEAF	-	-	-	-	-	-	SEAF/SEAM (3/6)
IND50.2	WM	WM	WF	WM	WM	WM	-	-	-	-	HISPM	WM	WM (6/8)
IND54	WM	WM	AIF	BM	HISPM	WM	-	-	-	-	-	-	WM (3/6)
IND55	HISPM	WM	WM	WM	WM	WM	-	-	-	-	-	-	WM (5/6)
IND56	AIM	HISPM	WM	WM	AIM	SEAM	-	-	-	-	-	-	Indeterminate

4.3 Non-Metric Analyses for Ancestry Estimation

    4.3.1 Optimized Summed Scoring Attributes (OSSA)

    OSSA scores (Hefner and Ousley 2014) are presented as their optimized summed score, ranging 0-6, in Table 4.14. Based upon score generated, predicted ancestry was either American Black or American White. Not all individuals had an OSSA score on account of taphonomic damage (n=16), and such individuals have "-" for a summed score and a predicted ancestry of "Indeterminate". If an individual had a score of "3" but had taphonomic damage, they were "Indeterminate" in ancestry. There were six cases where individuals had a score of "3" and two individuals with a "4". In cases where some, but not all, features were available for assessment, the mark "[a]" accompanies the number of features observed and marked by the observer. Only two individuals (IND52 and IND53) were estimated to be American Black.

Table 4.14: Optimized Summed Scores and Predicted Ancestry (Hefner and Ousley 2014).

Case Number	Optimized Summed Score	Predicted Ancestry
IND1	5	American White
IND2	3[a]	Indeterminate
IND3	4	American White
IND4	3[a]	Indeterminate
IND5	-	Indeterminate
IND6	6	American White
IND7	4[a]	American White
IND8	4	American White
IND9	3[a]	Indeterminate
IND10	4	American White
IND11	6	American White
IND12	-	Indeterminate
IND13	4[a]	American White
IND14	5	American White
IND15	3[a]	Indeterminate
IND16.1	-	Indeterminate
IND23	-	Indeterminate
IND28	-	Indeterminate
IND29	-	Indeterminate
IND34	-	Indeterminate
IND35	-	Indeterminate
IND37	3[a]	Indeterminate
IND38	4	American White
IND39	-	Indeterminate
IND45	4	American White
IND46	4	American White
IND47	5	American White
IND50.1	3[a]	Indeterminate
IND51	3[a]	Indeterminate
IND52	2[a]	American Black
IND53	2	American Black
IND54	6	American White
IND55	3[a]	Indeterminate
IND56	5	American White
IND57	4	American White

[a]One or more non-metric feature(s) damaged and unavailable for estimation.

4.3.2 Discriminant Functions for Two- and Three-Group Analyses

As stated in Chapter III, Hefner later expanded his OSSA scores method to become that of discriminant function analysis (DFA), with the inclusion of a seventh non-metric trait (nasal overgrowth) as well as a Hispanic reference population. This method employed the use of discriminant function analyses as opposed to converting summed scores to binary scores (Hefner 2015). Table 4.15 presents the estimation between American Blacks, American Whites and Hispanics with a three-way DFA. Using this method, more individuals (n=10) are estimated to be Hispanic, though this may simply be an outcome of having an additional population added – it is still possible that these individuals are of a different ancestry and did not identify as Hispanic in life. Table 4.16 gives the results of the two-way DFA. The two-way function predicted American White ancestry for all but one individual (IND53).

As in the sample used by Hefner and Ousley (2014), some individuals (n=11) were impossible to score given taphonomic damage. This resulted in only seven-trait and three-functions used, though Hefner (2015) does provide a five-trait function. Unfortunately, this five-trait function was not applicable for use in any of the individuals scored, as it required a score for nasal overgrowth, which was frequently damaged and unavailable for observation and scoring.

An estimation was possible for more individuals in this method than through the use of OSSA (Hefner and Ousley 2014). However, individuals estimated as "Indeterminate" through the use of OSSA scores did not necessarily translate to a Hispanic estimation with the discriminant function equations. Estimations of American White or American Black were still possible for these individuals, despite these being available ancestries using the OSSA method. In fact, there are instances (n=18) where individuals previously estimated as either American White or American Black with OSSA scores were estimated differently using Hefner (2015)

(Three-way DFA n=10; Two-way DFA n=1). American Black was the continued prediction for

IND53, and was predicted as such using the OSSA method (Hefner and Ousley 2014).

Table 4.15. Three-way DFA Scores and Predicted Ancestry (Hefner 2015).

Case Number	Trait-function	Cross-validation	DFA Score 1	DFA Score 2	Predicted Ancestry
IND1	Seven-trait	83.40%	3.011	1.69	Hispanic
IND2[a]	-	-	-	-	Indeterminate
IND3	Three-trait	63.10%	1.437	4.821	Am. White
IND4	Three-trait	63.10%	0.877	0.139	Am. White
IND5[a]	-	-	-	-	Indeterminate
IND6	Seven-trait	83.40%	2.834	1.824	Hispanic
IND7	Three-trait	63.10%	1.298	0.975	Am. White
IND8	Seven-trait	83.40%	0.47	1.663	Hispanic
IND9	-	-	-	-	Indeterminate
IND10	Seven-trait	83.40%	0.458	1.977	Hispanic
IND11	Seven-trait	83.40%	1.586	1.884	Hispanic
IND12	-	-	-	-	Indeterminate
IND13	Three-trait	63.10%	2.771	4.267	Am. White
IND14	Seven-trait	83.40%	1.284	1.781	Hispanic
IND15[a]	-	-	-	-	Indeterminate
IND16.1[a]	-	-	-	-	Indeterminate
IND23[a]	-	-	-	-	Indeterminate
IDN28[a]	-	-	-	-	Indeterminate
IND29[a]	-	-	-	-	Indeterminate
IND37	Three-trait	63.10%	1.051	2.176	Am. White
IND38	Three-trait	63.10%	2.561	3.851	Am. White
IND39[a]	-	-	-	-	Indeterminate
IND45	Seven-trait	83.40%	1.063	2.114	Hispanic
IND46	Seven-trait	83.40%	0.458	1.977	Hispanic
IND47	Three-trait	63.10%	1.929	2.411	Am. White
IND50.1	Three-trait	63.10%	2.807	2.646	Am. White
IND51[a]	-	-	-	-	Indeterminate
IND52	Three-trait	63.10%	0.21	0.416	Am. White
IND53	Seven-trait	83.40%	0.244	2.041	Hispanic/Am. Black
IND54	Seven-trait	83.40%	1.914	0.97	Hispanic
IND55	Three-trait	63.10%	3.227	3.478	Am. White
IND56	Three-trait	63.10%	3.895	-2.124	Hispanic
IND57	Three-trait	63.10%	2.104	4.544	Am. White

[a]One or more features unavailable for assessment – DFA not possible to run in these cases

Table 4.16. Two-way DFA Scores and Predicted Ancestry (Hefner 2015).

Case Number	Trait-function	Cross-validation	DFA Score	Predicted Ancestry
IND1	Seven-trait	86.00%	2.83	Am. White
IND2[a]	-	-	-	Indeterminate
IND3	Three-trait	85.40%	1.816	Am. White
IND4	Three-trait	85.40%	0.829	Am. White
IND5[a]	-	-	-	Indeterminate
IND6	Seven-trait	86.00%	3.084	Am. White
IND7	Three-trait	85.40%	1.338	Am. White
IND8	Seven-trait	86.00%	0.395	Am. White
IND9[a]	-	-	-	Indeterminate
IND10	Seven-trait	86.00%	0.415	Am. White
IND11	Seven-trait	86.00%	1.941	Am. White
IND12[a]	-	-	-	Indeterminate
IND13	Three-trait	85.40%	3.068	Am. White
IND14	Seven-trait	86.00%	1.739	Am. White
IND15[a]	-	-	-	Indeterminate
IND16.1[a]	-	-	-	Indeterminate
IND23[a]	-	-	-	Indeterminate
IND28[a]	-	-	-	Indeterminate
IND29[a]	-	-	-	Indeterminate
IND34[a]	-	-	-	Indeterminate
IND35[a]	-	-	-	Indeterminate
IND37	Three-trait	85.40%	1.118	Am. White
IND38	Three-trait	85.40%	2.865	Am. White
IND39[a]	-	-	-	Indeterminate
IND45	Seven-trait	86.00%	1.386	Am. White
IND46	Seven-trait	86.00%	0.415	Am. White
IND47	Three-trait	85.40%	2.05	Am. White
IND50.1	Three-trait	85.40%	2.982	Am. White
IND51[a]	-	-	-	Indeterminate
IND52	Three-trait	85.40%	0.203	Am. White
IND53	Seven-trait	86.00%	0.258	Am. Black
IND54	Seven-trait	86.00%	2.225	Am. White
IND55	Three-trait	85.40%	3.388	Am. White
IND56	Three-trait	86.00%	4.117	Am. White
IND57	Three-trait	85.40%	2.442	Am. White

[a]One or more features unavailable for assessment – DFA not possible to run in these cases

4.3.3 Expansion of Macromorphoscopic Traits

These data are available in Appendix B for future use in estimations for when a software program becomes available or for those able to apply the statistics necessary for estimations (e.g. k-nearest neighbor, naïve Bayesian, etc.) (Plemons and Hefner 2016).

4.4 Observer Error

Inter-observers were asked only to perform metric analyses on the individuals selected by the author (see Table 3.8). Observer error was analyzed in regards to how differing measurements impacted actual estimations depending on the metric method. Table 4.17 demonstrates the number of disagreements (i.e. when an inter-/intra-observer is above or below the original measurement by more than 2 mm) for each observer. The author chose to arbitrarily note when a measurement also exceeded 5 mm in difference from the original measurement to show measurements that were even more problematic. Also, within this table are the measurement names for each disagreement. Bolded entries account for instances where ancestry estimations changed from the original estimation made by the author. (Note: *All* measurements and side measurements taken are included in this table, not just the measurements selected by Stepwise Forward Wilks in FORDISC 3.1 (Jantz and Ousley 2010a). Mandibular measurements not employed by FORDISC 3.1 but used in [hu]MANid [Berg and Kenyhercz 2017] are also included.)

Table 4.17. Number of Disagreements between Observers with Measurement Names for Observer Error.

Case Number	# of Meas.	# and name of SLP Dis.		# and name of JHE Dis.		# and name of ARM Dis.		# and name of INTRA Dis.	
IND1	**23**	**2**	**XCB, EKB**	**2**	**GOL, EKB[b]**	-	-	**3**	**OBB, EKB, MDH**
IND2	26	-	-	-	-	0	N/A	0	N/A
IND3	27	-	-	-	-	0	N/A	2	EKB, DKB
IND4	26	-	-	-	-	0	N/A	1	DKB
IND6	41	1	GNI	1[b] (16)	MAL	-	-	2	XRH[c] (2x)
IND11	41	-	-	-	-	0[b] (14)	N/A	4	EKB, PAC, TLM23, XDA
IND12	**7**	2	FRC[c], PAC	**1**	**PAC**	-	-	1	FRC[c]
IND13	13	-	-	-	-	0	N/A	**2**	**FOL, MDH**
IND14	**24**	**3**	**OBB, MDH (2x)**	**2**	**BNL[c], MDH**	-	-	**1**	**DKB**
IND15	16	0	N/A	3	FRC, PAC, MDH	-	-	0	N/A
IND27	14	2	XRH, XDA	1	XDA	-	-	0	N/A
IND32	6	0[b] (1)	N/A	0[b] (1)	N/A	-	-	-	-
IND33	8	1	HML	0	N/A			1[b] (2)	HML
IND37	30	-	-	-	-	1[b] (5)	MDH	-	-
IND38	23	-	-	-	-	-	-	1	MAB
IND39	12	-	-	-	-	0	N/A	-	-
IND40.1	12	1	XRH[c]	1	XRH[c]	-	-	1	XRH[c]
IND40.2	10	1	XRH[c]	1	XRH	-	-	1[b] (1)	XRH[c]
IND40.3	7	1	TIBNFX	0	N/A			0[b] (1)	N/A
IND46	26	-	-	-	-	0	N/A	2	XCB, MDH
IND47[a]	**34**	-	-	-	-	0[b] (20)	N/A	**1**	**MDH**
IND50.1	12	-	-	-	-	-	-	2	OCC, MDH
IND50.2	14	-	-	-	-	-	-	0	N/A
IND51	14	0	N/A	1	MAL	-	-	2	MAL, UFHT
IND52	25	-	-	-	-	0	N/A	3	BPL, MAL, MDH
IND53	23	-	-	-	-	0	N/A	0	N/A
IND57	23	-	-	-	-	0	N/A	1	MDH

[a]Postcranial measurements included.

[b]Not all measurements taken – number is specified.

[c]Cases where observer measurements above or below original measurement by more than 5 mm.

4.4.1 Error with Cranial Measurements

Here I present the measurements that had the highest levels of disagreement among observers, along with an indication of whether the measurements appeared to have an impact on the ancestry estimation of that individual. Interpretations of why certain measurements were problematic and their potential importance in ancestry estimations are provided in greater detail in Chapter V.

The measurement "Mastoid Height" (MDH) was the measurement that had the largest amount of disagreement between observers, with 12 cases of disagreement, though none of these disagreements exceeded 5 mm in difference. For the 10 individuals where MDH was used by FORDISC 3.1 (Jantz and Ousley 2010a), five individuals had changes in their ancestry estimations from the original estimation done by the author. However, observers predicted a total of eight different ancestries for these five individuals that had a disagreement in the MDH measurement.

Other cranial measurements also proved challenging, with biorbital breadth (EKB) having the second-largest amount of disagreements in this study (n=5). One measurement was greater than 5 mm in distance from the original measurement. This is significant as it implies that there may be other factors occurring that induced such an error. Three cases of disagreement for this measurement occurred for IND1 with none of the observers (i.e. SLP, JHE and HBD) agreeing with the first measurement. FORDISC 3.1 (Jantz and Ousley 2010a) estimations that used EKB changed from the original ancestry estimation three out of five times.

There were four cases of disagreement with the palate length measurement (MAL); however none exceeded 5 mm in difference. There were no cases where disagreements between

observers in MAL values changed the ancestry estimations predicted by FORDISC 3.1 (Jantz and Ousley 2010a).

Lastly, there were four disagreements for the parietal chord (PAC) with none being greater than 5 mm from the original measurement. There is one case (IND12) out of the four FORDISC 3.1 runs where PAC may have had an impact in changing the ancestry estimation from the original estimation. SLP and HBD also record large differences for the frontal chord measurement (FRC) for this individual. Given that JHE did not have a disagreement from the original FRC and still resulted in a different ancestry prediction, FRC likely was not what altered the estimation.

4.4.2 Error with Mandibular Measurements

The maximum ramus height (XRH) measurement was the only measurement that had a large number of disagreements (n=9), with seven of the nine having more than a 5 mm difference from the original measurement. For three individuals measured for observer error (IND6, IND40.1, IND40.2), IND40.1 has different ancestry estimations between the observers. The original estimation for this individual, as seen in Table 4.16, was a Black Female (BF); however SLP, JHE and HBD's later estimations with secondary measurements indicated the individual to be possibly American Indian Female (AIF). The large discrepancies in XRH measurements for all observers did not alter ancestry estimations for the two other individuals (IND6 and IND40.2).

4.5 Comparison of Ancestry Estimations for All Methods

Below (Table 4.18) are the ancestry estimations predicted by each method compiled for each individual. In cases where more than one ancestry was predicted (e.g. FORDISC 3.1 and [hu]MANid), the most common/frequent ancestry estimated by all observers and all data types

(i.e. metric, morphoscopic, morphometroscopic) was used as the final estimation. The author weighed this final estimation as one against all other estimations from other methods to make comparison between each method equal.

Individuals in Table 4.18 that are bolded demonstrate cases where more than half of the methods used provide the same estimation. The author did not count or bold cases where "Indeterminate" estimations had the highest frequency for an individual. Indeterminacy in these cases was due to taphonomic damage or disagreement of ancestry between observers, thereby making the method not applicable for the unknown. If an individual's ancestry estimation came from using only one method, the estimation was not bolded as there was no comparison to make to other methods.

Table 4.18. Overall Predicted Ancestries for All Methods Used (Berg and Kenyhercz 2017; Hefner and Ousley 2014; Hefner 2015; Jantz and Ousley 2010a) with Most Common Ancestries

Case Number	FORDISC 3.1 Cranium	FORDISC 3.1 Mandible	FORDISC 3.1 Postcrania	OSSA	Hefner (2015) 3-way DFA	Hefner (2015) 2-way DFA	(hu)MANid
IND1	Indeterminate	-	-	American White	Hispanic	American White	-
IND2	JF	-	-	Indeterminate	Indeterminate	Indeterminate	-
IND3	AM[c]	-	-	**American White**	**American White**	**American White**	-
IND4	AF	-	-	Indeterminate	American White	American White	-
IND5	BF	-	-	Indeterminate	Indeterminate	Indeterminate	-
IND6	**WM**	GTM	**WM**	**American White**	Hispanic	**American White**	**WM**
IND7	BF	-	-	**American White**	**American White**	**American White**	-
IND8	HF	WF	-	American White	Hispanic	American White	BF
IND9	BF	-	-	Indeterminate	Indeterminate	Indeterminate	-
IND10	**WM**	GTM	-	**American White**	Hispanic	**American White**	**WM**
IND11	AM[c]	Indeterminate	-	American White	Hispanic	American White	Indeterminate
IND12	WM	-	-	Indeterminate	Indeterminate	Indeterminate	-
IND13	AM	-	-	**American White**	**American White**	**American White**	-
IND14	VM	-	-	American White	Hispanic	American White	-
IND15	AF	-	-	Indeterminate	Indeterminate	Indeterminate	-
IND16.1	Indeterminate	HM	WM	Indeterminate	Indeterminate	Indeterminate	Indeterminate
IND16.2	-	-	BF	-	-	-	-
IND23	Indeterminate	JM	BM	Indeterminate	Indeterminate	Indeterminate	HISPM
IND25[a]	-	JF	-	-	-	-	AIF
IND26	-	GTM	-	-	-	-	BM
IND27	-	BM	-	-	-	-	HISPM
IND28	Indeterminate	Indeterminate	-	Indeterminate	Indeterminate	Indeterminate	BF
IND29	Indeterminate	GTM	-	Indeterminate	Indeterminate	Indeterminate	WM
IND30	-	-	-	-	-	-	BM
IND31	-	Indeterminate	-	-	-	-	AIF
IND32	-	Indeterminate	-	-	-	-	BF
IND33	-	WM	-	-	-	-	HISPM

Case Number	FORDISC 3.1 Cranium	FORDISC 3.1 Mandible	FORDISC 3.1 Postcrania	OSSA	Hefner (2015) 3-way DFA	Hefner (2015) 2-way DFA	(hu)MANid
IND34	Indeterminate	JM	-	Indeterminate	Indeterminate	Indeterminate	BF
IND35	Indeterminate	**BF**	-	Indeterminate	Indeterminate	Indeterminate	**BF**
IND36	-	**-**	-	-	-	-	WM
IND37[a]	AM	Indeterminate	-	Indeterminate	American White	American White	SEAM
IND38	HM	-	-	**American White**	**American White**	**American White**	-
IND39	AF	-	-	Indeterminate	Indeterminate	Indeterminate	-
IND40.1	-	WF	-	-	-	-	AIF
IND40.2[a]	-	CHM	-	-	-	-	AIM
IND45	BF	-	-	American White	Hispanic	American White	-
IND46	**WM**	-	-	**American White**	Hispanic	**American White**	-
IND47	AF	Indeterminate	BF	American White	American White	American White	AIF
IND48	-	**BF**	-	-	-	-	**BF**
IND49[b]	-	JM	-	-	-	-	SEAF/SEAM
IND50.1	BF	-	-	Indeterminate	American White	American White	-
IND50.2	-	Indeterminate	-	-	-	-	WM
IND51	JF	-	-	Indeterminate	Indeterminate	Indeterminate	-
IND52	AF	-	-	American Black	American White	American White	-
IND53	AF	-	-	**American Black**	Hispanic/**American Black**	**American Black**	-
IND54	**WM**	JM	-	**American White**	Hispanic	**American White**	**WM**
IND55	BF	**WM**	Indeterminate	Indeterminate	**American White**	**American White**	**WM**
IND56[a/b]	AM[c]	GTM	-	American White	Hispanic	American White	Indeterminate
IND57	GTM	-	-	**American White**	**American White**	**American White**	-

[a]Where Native American (e.g. AF/AIF/AM/AIM) is predicted in one method and estimated as either Hispanic or Asian in another.
[b]Where two different Asian groups (e.g. CHM, JM, SEAM, etc.) or two Hispanic groups (e.g. GTM, HM, HISPM, etc.) are predicted by different methods
[c]Estimation of American Indian given by context, despite low typicalities calculated by FORDISC 3.1 (Jantz and Ousley 2010a).

In Table 4.18, there is a wide variety of predictions where very few seem to be in a majority agreement. Only 13 (bolded) of the total 49 cases (~27%) have instances where more than half of the methods applied to them agree.

FORDISC 3.1 and (hu)MANid did use approximately the same number of population groups to generate estimations. However, possibly due to more measurements available for use through (hu)MANid, there were many instances where (hu)MANid was able to provide more physical generated estimations than FORDISC 3.1. The GUI requires a minimum of two measurements inputted into program to provide an estimation, unlike FORDISC 3.1 which requires at least four measurements. For the 28 mandibles used in this study, (hu)MANid generated an estimation 100% of the time while FORDISC 3.1 predicted the ancestry in only 22 cases (~79%) due to taphonomic damage impacting the number of available measurements. In this study, a maximum of six measurements were used to estimate ancestry in FORDISC 3.1 (Table 4.6) while (hu)MANid had a maximum of nine, thus perhaps capturing more variation between mandibles and the populations used in this method.

Comparison of Hefner and Ousley (2014) and Hefner (2015). In Table 4.18, OSSA scores and the two discriminant function scores show that 22 of 35 cases (~63%) (not including "Indeterminate" estimates) had at least two of these estimates that were in agreement with one another. Estimations for all of these individuals were American White except IND53, estimated to be American Black. Logistically, there could not be two cases of agreement for the estimation of Hispanic as the two-way discriminant function and the OSSA scores use only American Whites and American Blacks as their reference.

Almost half of these estimates (n=10) however, came from agreement between the OSSA method and the 2-way discriminant function where the American White and American Black

groups are based off of the same reference collections. In only seven of the 22 total cases

(~32%) there was agreement between all three methods.

     Comparison of (hu)MANid to Other Non-Metric Methods. Table 4.18 outlines cases

where at least three of the four non-metric methods (OSSA scores [Hefner and Ousley 2014],

Three-Way DFA [Hefner 2015], Two-Way DFA [Hefner 2015] and [hu]MANid [Berg and

Kenyhercz 2017]) agreed. There was agreement in only four out of 15 individual cases (~27%)

(again, not counting an agreement between "Indeterminate" scores).

     Comparison of FORDISC 3.1 to Non-Metric Methods. There were 35 possible

comparisons between FORDISC 3.1 for the cranium and all of the non-metric methods. Within

these 35, there were 11 cases (~31%) where three or more of the methods yielded the same

result. Again, the author did not count estimates that were "Indeterminate" as an agreement.

Seven of these 11 agreements originated from agreement only between the non-metric cranial

methods.

     In cases where all skeletal elements (i.e. cranium, mandible, postcranium), were entered

into FORDISC 3.1 (n=5), agreement between at least two elements within the program occurred

only once, with the individual (IND6) being classified as a White Male (WM). This individual

was classified as a WM with the OSSA method (Hefner and Ousley 2014), the two-way

discriminant function (Hefner 2015) and with (hu)MANid (Berg and Kenyhercz 2017). This

individual has the highest amount of agreement between all methods, with five of the seven

methods agreeing.

     Comparison of (hu)MANid to FORDISC 3.1. There were 26 cases where mandibular

measurements from FORDISC 3.1 were compared against estimates given by (hu)MANid (Table

4.18). Not counting the estimates that were "Indeterminate" by both programs, the two agreed

three times (~12%). In two of these cases, the individual was estimated to be a Black Female (BF) and in one a White Male (WM). However, there was sex agreement in 14 of the 26 cases (~54%).

IND25 and IND40.2 yielded estimates from these two methods that were not identical. FORDISC 3.1 estimates both of these individuals to be of Asian ancestry while (hu)MANid estimates them both to be of Native American descent. Four of the 26 comparable estimations given by (hu)MANid do happen to corroborate with cranial estimations given by FORDISC 3.1 (~15%), yet none of the accompanying FORDISC 3.1 mandibular estimations match with either.

4.6 Summary

As observed in Table 4.18, only 13 out of 49 individuals (~27%) had ancestry estimates that agreed based on the methods applied. Implications of this will be discussed in Chapter V, with specific regard to NAGPRA implications in section 5.5.2. OSSA scores and discriminant function analyses applied to cranial non-metric traits agreed the most in 22 out of 35 cases. The analysis of mandibular metrics between (hu)MANid and FORDISC 3.1 however never agreed. Estimates were varied depending on the populations used in each of the methods as a reference, with FORDISC 3.1 and (hu)MANid predictably having the greatest variability in estimations.

Given the breadth of variation of these ancestry estimations predicted by these methods, a forensic anthropologist may feel cause for concern. Chapter V expands upon the implications of these differences and disagreements with potential explanations for these incongruities. This chapter also discusses observer error and its impact on this research in greater detail. Lastly, the author addresses future avenues of research and possible solutions for the incomparability of ancestry estimation methods.

Chapter V:

Discussions and Conclusion

5.1 Introduction

In this final chapter, results for the research questions outlined in Chapter I are revisited

and the potential value of the collections, such those from the ACCO, are discussed. Implications

for this study, observer error and limiting factors follow. Concerns for ancestry estimation

method comparability are addressed with limitations and issues. Lastly, future directions

concerning the creation and development of methods used for the biological profile are provided.

5.2 Research Questions

5.2.1 (Q1) Are metric and non-metric methods specifically (e.g. Berg 2015; Hefner 2015;

Hefner and Ousley 2014; Jantz and Ousley 2010a) for estimating ancestry comparable (i.e. do

they generate the same ancestry estimation)? How do methods with similar reference

populations influence comparability?

Comparison of Hefner and Ousley (2014) and Hefner (2015). These non-metric methods

had many instances of agreement between one another (22 out of 35/~63%). Two of the groups

(American White and American Black) derive from the same reference collections for both

methods, which possibly induced some cases of agreement. Furthermore, these methods also use

essentially the same morphoscopic features, only applying them in different ways.

Though there are these similarities between these methods, as stated in Chapter IV, there

were only seven out of 22 cases of total agreement (~32%). This is likely due to the addition of

the Hispanic group within the Three-Way DFA, a group that neither the OSSA method (Hefner

and Ousley 2014) nor the Two-Way DFA method (Hefner 2015) employs. The addition of the

seventh non-metric trait in the DFA method (nasal overgrowth/NO) may have had an influence on these changes, but this is not verified.

The issue of the Hispanic group in ancestry classification is well-known in forensic anthropology (Spradley et al. 2008), and while their inclusion in the creation of the method is a step to increase recognition of human variation, there are still problems. Currently the term "Hispanic" largely relates to individuals who are Spanish-speaking and identify with the specific culture. Furthermore, such individuals could be from any location where Spanish is the predominant language. As this accounts for hundreds of millions of people, using this term is overly general and cannot appropriately describe the individuals in collections such as those from the Pima County Office of the Medical Examiner (PCOME) – the collection used for the creation of this method. As such, while this method does increase variation, such limitations and considerations when applying this term must be acknowledged.

Comparison of (hu)MANid to Other Non-Metric Methods. Berg's (2015) and Berg and Kenyhercz's (2017) data can also be compared against all other non-metric methods. Though there are more composite and individual/specific groups used in this method, the US White and US Black samples are both from the same collection(s) as the methods used by Hefner (2015) and Hefner and Ousley (2014). The Hispanic sample, however, is quite different – the composite group is made up by 89 Guatemalan/Mayan males as opposed to only 30 males identified by the PCOME. The females of the Hispanic sample (n=14) are also only comprised of Guatemalans/Mayans in (hu)MANid.

The addition of other population groups likely affected the lack of agreement between this method and the other three non-metric methods (Hefner and Ousley 2014 and Hefner 2015). (hu)MANid is capable of recognizing a higher degree of variation between these other

populations and also uses features only identified on the mandible. As the other non-metric

methods used in this thesis had specific focus on the cranium, it is possible that some of the

discrepancy arose from the difference in skeletal elements observed. This is cause for concern as

the cranium and the mandible are not always collected together in forensic or bioarchaeological

settings.

Comparison of FORDISC 3.1 to Non-Metric Methods. Comparing non-metric methods to

FORDISC 3.1 proves challenging for many of the same reasons found in comparing them to

Berg (2015) and Berg and Kenyhercz (2017). There are several more populations considered

using FORDISC 3.1 as opposed to a maximum of three populations used in the methods

developed by Hefner.

FORDISC 3.1 (Jantz and Ousley 2010a), OSSA scores (Hefner and Ousley 2014) and

discriminant function analyses for morphoscopic traits (Hefner 2015) do all use the cranium and

as such, application of all methods is possible. However, the greater number of populations used

by FORDISC 3.1 could potentially make forensic anthropologists more confident in metric data

as opposed to morphoscopic data. The future development of a program or discriminant function

for additional morphoscopic traits with more population groups utilized (Hefner 2018) could

help to balance metric data and morphoscopic data better.

(Q1a) How do software programs (e.g. FORDISC 3.1 [Jantz and Ousley 2010a], and

(hu)MANid [Berg and Kenyhercz 2017]) compare against one another? The reference groups

between (hu)MANid and FORDISC 3.1 are more comparable for the mandibular measurements.

(hu)MANid also includes more reference population groups on a global level. Much of the

Forensic Data Bank (FDB) derives from the William M. Bass Collection, though current forensic

anthropologists from around the United States can contribute their data as well. Locations and

time periods in FORDISC 3.1 are much more specific, though there is often a counterpart to be found in the composite sample used by Berg and Kenyhercz (2017) in (hu)MANid. For example, no Japanese samples are included in (hu)MANid while they are present in the FDB; however, the (hu)MANid databank includes Cambodians, Thais, and Koreans in their pooled samples while FORDISC 3.1 does not.

The low corroboration, as noted in Chapter IV, between these two methods is striking, as it has been described previously that the two are actually quite comparable and yielded even "identical" results in some cases (Berg and Kenyherz 2017:1595). Contrasting estimates given between the different programs are possibly a result of HBD using the most common/frequent ancestry estimations between observers (e.g. Table 4.13) and between methods (e.g. Table 4.18) for each of the individuals instead of accounting for other estimations that could have high posterior probabilities and appropriate typicalities as well. The author (HBD) used the most common/frequent ancestry estimation for the reason that independent of other methods, these would be the final estimations that the author would provide in an official report.

Other reasons for the lack of corroboration may relate to more population groups being accessible when only observing the cranium through FORDISC 3.1, particularly those of American Indian ancestry as was seen in one of the four cases. The other three, however, classified the individuals as White Males (WM) which is a useable reference group when only analyzing mandibular measurements into FORDISC 3.1.

How do they Compare? Taken all together, these seven methods produce few instances of agreement, and many of these methods do not appear to be comparable, aside from methods developed or co-developed by Hefner (Hefner and Ousley 2014; Hefner 2015). Only in respect to methods developed by Hefner can the author *tentatively* say that there is comparability. Stated

in Chapter IV, only 13 of the total 49 cases (~27%) have instances where more than half of the methods applied to them agree. This is more than likely due to the variability found in the reference collections used to create these methods. Even with some populations being presented as the same (e.g. Vietnamese Males in FORDISC 3.1 and [hu]MANid), the individuals referenced are not congruent.

Furthermore, the disparities in sample sizes for these methods also prove problematic, as most instances of agreement between methods only occurred with Black Females, Black Males, and White Males. The reference collections used between these particular groups is often the same for each of the methods, perhaps contributing to a higher level of agreement.

Other sample sizes (e.g. American Indians, Hispanics and various Asian populations) remain small in comparison. This is a challenge as many of the samples are also specific to only one group of people, thus preventing total agreement between methods. For instance, though (hu)MANid has a higher number of American Indians in its reference database, this entire group is a composite of two cultures: the Arikara in South Dakota, dating from 1679-1733 C.E. and the Hohokam in Central Arizona which dates from 1150-1450 C.E. (Berg 2015:48). FORDISC 3.1's reference for this group is from the mid-late 19th century (Jantz and Ousley 2010b:15), thus being slightly closer to modern temporal significance; however, due to secular changes for some populations occurring even within a 50 year period (Bogin 1999), these groups today likely do not resemble recent generations. This sample from FORDISC 3.1 is also specific to the American Southwest and does not begin to capture the breadth of variation within this group.

Problems such as these created a great deal of discrepancy between the methods used, with more population groups used in some methods creating more differences in estimations. As many as three different estimations could be seen for an individual alone (n=13) based on Table

4.18. Some of these particular individuals and others with just two different estimations derive from populations that likely have a high relatedness to one another, thus creating this confusion.

IND49 is an example of such possible relatedness, being estimated as a Japanese Male (JM) by FORDISC 3.1 and as a Southeastern Asian Female/Male (SEAF/SEAM) by (hu)MANid. As (hu)MANid does not use a Japanese reference group, the individual can be estimated as broadly Asian. It is interesting that the GUI did not estimate Northeastern Asian Female/Male (NEAF/NEAM), however, as this group is closer in relation to the Japanese group, or that FORDISC 3.1 did not estimate Vietnamese Male (VM) for the same reason. It is possible that the greater reference sample size of SEAM (315 individuals) over the sample size of NEAM (199 individuals) (Berg and Kenyhercz 2017) or VM (51 individuals) in FORDISC 3.1 (Jantz and Ousley 2010b:16) caused this, as more variation is likely recognized.

Populations with likely relatedness (Getz, personal communication, 2019) could be those of an Asian ancestry (e.g. Japanese and Southeastern Asian as seen with IND49), Hispanic ancestry (individuals from the Pima County Office of Medical Examiner and Guatemalan as seen in IND10), or more broadly exemplifying possible relatedness between American Indian groups and Asian or Hispanic groups (as seen with IND25 and IND56 respectively).

5.2.2 (Q2) If methods generate different estimations, what are the aspects of these methods that researchers need to consider? What improvements would elevate the accessibility and agreement between these methods? Can these methods be used to complement each other?

Given the wide variety of outputs generated through the use of all of these methods, it is advised that more awareness should be placed on the reference populations available for classification when using these methods. Though there appears to be more agreement between the methods developed by Hefner (Hefner 2015; Hefner and Ousley 2014), it is necessary to

recognize that these three methods (OSSA scores, Three-way and Two-way Discriminant Function Analyses) offer little difference in reference populations. The addition of the Hispanic group in the Three-way DFA allows for more human variation to be observed, though it is only through seven morphoscopic traits that we are seeing this.

Non-metric methods are continuing to undergo development (e.g. Hefner and Linde 2018) and it is the hope that these will eventually have their own software where statistical analyses can be run by forensic anthropologists to estimate ancestry (Hefner 2018). Until that time, the forensic anthropologist must look for ways for these methods (or others) to complement one another so as to improve the accuracy of classifications and quicken positive identifications.

The use of non-metric methods should certainly continue, though perhaps as a starting point to other metric methods, including those not evaluated in this thesis and ones that examine different skeletal elements. The metric methods used here can corroborate such ancestry estimations as the reference populations used by the metric methods in this study are more diverse. Should these methods continue to disagree, a study of the context of the skeletal remains upon excavation may prove helpful, where the demography may narrow/expand possibilities.

However, provenience and context may not always be available, as was the case for the collection housed at the ACCO. Except in the cases where populations identified are well-established in reference collections (e.g. American Blacks and American Whites), these methods did not agree. The metric methods used do, however, have the potential to complement one another. Even in cases where the methods did not agree entirely, there were cases where agreement on a general geographic area were possible. Though this does not definitively answer the question of ancestry for such individuals, it still indicates to the anthropologist several things: (1) the population of the unknown individual may not be included in the reference collections;

(2) sample sizes of the possible population may be too small to properly recognize variation; (3)

and the user can know to some degree to which groups an unknown individual does not belong.

When using software, the user should become aware of the patterns that arise through

each run of the program. While the classification may not point directly to a specific ancestry,

patterns of measurements used, sexes used, distances, typicalities, etc. can indicate a general

understanding of the unknown individual.

Examples of FORDISC 3.1 Patterns being used to Make Estimations. IND11 posed an

interesting challenge when being evaluated through FORDISC 3.1 where recognizing patterns

was crucial. Upon the first run (done with 18 cranial measurements) where all populations were

included, all female groups and the Guatemalan male group resulted in F typicalities that were all

< 0.01 (Figure 5.1). The second run removed all of these groups. Though there were five other

groups that had typicalities less than 0.05, these remained in the event that the female groups and

the Guatemalan male group were making an impact on the original classification. Figure 5.1

shows the results from the first run and Figure 5.2 shows the results from the second.

```
---------------------------------------------------------------------------
  Group      Classified     Distance          Probabilities
              into           from      Posterior   Typ F   Typ Chi   Typ R
---------------------------------------------------------------------------

    VM         **VM**          24.7       0.758     0.160    0.133    0.184 (40/49)
   CHM                         27.7       0.171     0.081    0.067    0.014 (73/74)
    JM                         31.4       0.027     0.031    0.026    0.016 (181/184)
    HM                         31.8       0.022     0.028    0.023    0.109 (155/174)
    WM                         33.6       0.009     0.017    0.014    0.038 (277/288)
    AM                         34.4       0.006     0.016    0.011    0.135 (45/52)
    BM                         35.8       0.003     0.010    0.008    0.077 (96/104)
   GTM                         36.5       0.002     0.009    0.006    0.015 (67/68)
    AF                         38.1       0.001     0.007    0.004    0.034 (28/29)
    JF                         38.3       0.001     0.005    0.004    0.009 (113/114)
    HF                         42.2       0.000     0.002    0.001    0.026 (37/38)
    WF                         45.1       0.000     0.001    0.000    0.006 (175/176)
    BF                         50.2       0.000     0.000    0.000    0.014 (68/69)
---------------------------------------------------------------------------
  Current Case is closest to VMs
---------------------------------------------------------------------------
```

Figure 5.1. First Analysis of IND11 using FORDISC 3.1 (Jantz and Ousley 2010a). Red signifies cases where measurements are extremely atypical (typicality <0.05) for that particular group.

```
-------------------------------------------------------------------------------
  Group       Classified      Distance          Probabilities
                into            from      Posterior   Typ F   Typ Chi   Typ R
-------------------------------------------------------------------------------


     VM                          39.7         0.546    0.010    0.005    0.020  (48/49)
    CHM                          41.6         0.208    0.005    0.003    0.014  (73/74)
     HM                          42.9         0.113    0.004    0.002    0.034  (168/174)
     JM                          43.2         0.095    0.003    0.002    0.005  (183/184)
     AM                          46.4         0.019    0.002    0.001    0.038  (50/52)
     BM                          47.3         0.012    0.001    0.001    0.019  (101/103)
     WM                          48.7         0.006    0.001    0.000    0.003  (286/287)
-------------------------------------------------------------------------------
  Current Case is too dissimilar to all groups; all TPs <  0.01
-------------------------------------------------------------------------------
```

Figure 5.2. Second Analysis of IND11 using FORDISC 3.1 (Jantz and Ousley 2010a). Red signifies cases where measurements are extremely atypical (typicality <0.05) for that particular group.

The second run (done with 20 cranial measurements) yielded results that indicated that the individual was "too dissimilar to all groups" as none of the typicalities were deemed acceptable by the program. This pattern is typically seen when an individual is not from one of the reference groups included in the analysis. Though this individual was previously estimated to be of European descent by the ACCO, the author hypothesizes that this individual could possibly be of Native American ancestry. It is also possible that there is little representation of the group that this individual would most closely relate to in FORDISC 3.1. This may be the case also for IND3 and IND37. This possibility still corresponds with the hypothesis that this individual is of Native ancestry as the Native groups that are counted in FORDISC 3.1 include only those found in the Southwest region of the United States (Jantz and Ousley 2010b:15). Moreover, the sample size of this group between the sexes are also among the smallest samples in FORDISC 3.1.

Similar results to those described above for IND11 were common in FORDISC 3.1 for individuals who were likely of Native American ancestry. The recognition of such patterns in the program was necessary for many of these estimations. Such results did not occur for every individual possibly of Native ancestry, however, as was the case for IND53. In each run of the

program, the individual was classified into the Amer-Indian Female group as can be seen in

Figures 5.3 and 5.4, despite other groups having low typicalities. The author removed female

groups other than AF in the second run as all were in the bottom half of likelihood and had

posterior probabilities that equaled zero. Chinese and Japanese Males were also removed from

the second run as it was clear that the individual was likely not of Asian ancestry.

```
-------------------------------------------------------------------------
  Group      Classified    Distance         Probabilities
              into          from     Posterior  Typ F  Typ Chi   Typ R
-------------------------------------------------------------------------

   AF        **AF**          26.3      0.597    0.092   0.068   0.133 (26/30)
   AM                        27.2      0.382    0.068   0.055   0.222 (42/54)
   HM                        35.2      0.007    0.007   0.006   0.042 (204/213)
   GTM                       35.2      0.007    0.008   0.006   0.013 (76/77)
   BM                        36.9      0.003    0.004   0.003   0.017 (119/121)
   JM                        37.5      0.002    0.004   0.003   0.005 (195/196)
   CHM                       38.2      0.002    0.003   0.002   0.013 (74/75)
   JF                        44.2      0.000    0.000   0.000   0.008 (124/125)
   WM                        46.1      0.000    0.000   0.000   0.002 (473/474)
   HF                        46.1      0.000    0.000   0.000   0.021 (47/48)
   BF                        46.8      0.000    0.000   0.000   0.013 (74/75)
   VM                        47.1      0.000    0.000   0.000   0.020 (48/49)
   WF                        56.1      0.000    0.000   0.000   0.004 (257/258)
-------------------------------------------------------------------------
  Current Case is closest to AFs
-------------------------------------------------------------------------
```

Figure 5.3. First Analysis of IND53 using FORDISC 3.1 (Jantz and Ousley 2010a). Red signifies cases where measurements are extremely atypical (typicality <0.05) for that particular group.

```
-------------------------------------------------------------------------
  Group      Classified    Distance         Probabilities
              into          from     Posterior  Typ F  Typ Chi   Typ R
-------------------------------------------------------------------------

   AF        **AF**          22.7      0.562    0.179   0.121   0.036 (27/28)
   AM                        23.3      0.428    0.147   0.107   0.245 (40/53)
   HM                        32.4      0.005    0.015   0.009   0.025 (197/202)
   BM                        33.1      0.003    0.013   0.007   0.038 (101/105)
   GTM                       33.5      0.003    0.012   0.006   0.013 (76/77)
-------------------------------------------------------------------------
  Current Case is closest to AFs
-------------------------------------------------------------------------
```

Figure 5.4. Second Analysis of IND53 using FORDISC 3.1 (Jantz and Ousley 2010a). Red signifies cases where measurements are extremely atypical (typicality <0.05) for that particular group.

Both IND11 and IND53 underwent inter-observer and intra-observer error evaluation

with Dr. Michael observing. Estimations generated through subsequent runs with Dr. Michael's

measurements and the observer's second measurements are equal to those of the original analyses.

5.3 The Value of ACCO as a Manufactured Collection

Though the sample used for this study eliminated certain age cohorts (i.e. those definitively less than 17 yrs.), a chi-squared test was done based off of information in Table 3.2 to assess the significance of the male to female ratio to demonstrate some of Komar and Grivas' (2008) argument. The author chose an alpha level of 0.05. Males are significantly overrepresented within the collection ($p < 0.05$) if it is expected that there should be approximately 24.5 (50%) males and 24.5 (50%) females present. This is based on current data presented by the US Census Bureau (US Census Bureau 2017:Table S0101).

Other assessments (i.e. cause and manner of death and donation bias) done by Komar and Grivas (2008) were not able to be evaluated through this study as much is unknown regarding the provenience of many of the individuals. Looking at Table 4.18, however, many ancestry estimations of these individuals are not common to the state of Idaho's demographic (e.g. Native American population in Idaho is approximately 1.29% of the total population [US Census Bureau 2017:Table B03002] which is overrepresented in the ACCO).

More can be learned from collections such as the one found at ACCO even though they are not representative. This thesis indicates that testing methods pertaining to the biological profile in forensic anthropology for their comparability as well as their efficacy is possible and essential. Though there was one positively identified individual in the collection to test efficacy and accuracy, it was possible to test method comparability in other aspects of the profile.

Komar and Grivas (2008) investigated the demographic in documented skeletal collections (i.e. Maxwell Museum Documented Collection), giving specific notice to donation

bias and cause/manner of death along with other components of the biological profile. They also observed demographics within a medico-legal/forensic sample, similar to the collection at ACCO. Such collections consist of a very different type of demographic. These offices obtain remains through a largely unconventional route in some cases (e.g. remains are unclaimed by family) or in others, they persist in offices pending positive identifications for forensic and archaeological matters (e.g. Native American Graves Protection and Repatriation Act of 1990).

5.4 Observer Error

   5.4.1 Error with Cranial Measurements

   Cranial measurements have been evaluated in the literature for inter-/intra-observer error, as definitions and the uses for these measurements have changed over time and resources (e.g. *Standards for Data Collection from Human Skeletal Remains* [Buikstra and Ubelaker 1994]) used for obtaining definitions are also varied. All observers who participated in this study used *Standards* (Buikstra and Ubelaker 1994) to take their cranial measurements. Other options include the *Data Collection Procedures 2.0* (DCP 2.0) which is free for use online (Langley et al. 2016).

   Langley et al. (2018) used the *DCP 2.0* to evaluate observer error in cranial, mandibular, and postcranial elements. There were four observers total at varying levels of experience and approximately 200 measurements taken by each. The authors found that for cranial measurements, interorbital breadth (DKB), mastoid height by sight (MDH$_{SIGHT}$) (now omitted from *DCP 2.0*) and mastoid height by tip (MDH$_{TIP}$) had the highest levels of both inter-observer and intra-observer error.

   Though the author used Buikstra and Ubelaker (1994) for instructions on how to take cranial measurements, the author recommends using the *DCP 2.0* (Langley et al. 2016) at least

for MDH, as there are clearer cranial landmarks referenced. The authors define MDH$_{TIP}$ in the

DCP 2.0 as "the direct distance between porion and mastoidale" (Langley et al. 2016:69).

Definitions of these cranial landmarks are elsewhere in the *DCP 2.0.* Buikstra and Ubelaker

(1994), however, define this measurement as the "vertical projection of the mastoid process

below and perpendicular to the eye-ear (Frankfort) plane" (Buikstra and Ubelaker 1994:77).

Each observer had at least one instance of disagreement with this measurement (MDH)

and such disagreement matches what Langley et al. (2018) discussed. Confusion with how to

properly take the measurement may have caused these disagreements. If there is a correlation

between these disagreements and the number of cases where the ancestry changed (n=8), it is

possible that a large enough weight is put on this measurement for discrepancies in estimations

to occur.

For the measurement with the second-highest degree of disagreements in the present

study, biorbital breadth (EKB), identifying the cranial landmarks used for this measurement may

have been challenging. This could be due to potential difficulty in finding the correct point where

the orbit is divided into equal upper and lower halves (Buikstra and Ubelaker 1994:71). It is also

possible that the first measurement taken was done so incorrectly, only to be caught by the

observers. There were changing ancestry estimations in three out of the five disagreements for

EKB, however, all three of these also had disagreements for the MDH measurements. Therefore

it is possible that while EKB may have contributed to the change in ancestry estimation, it may

not have had the greatest impact.

For the measurement palate length (MAL), the observer is required to know the location

of the cranial landmark prosthion, which was frequently impacted by taphonomic damage and

was not taken in all cases along with possible bony growth on alveolon which could impact the

placement of the calipers. None of the disagreements for MAL resulted in changes from the original ancestry estimation.

As stated in Chapter IV, there were four disagreements for the parietal chord measurement (PAC). As this measurement requires recognizing cranial landmarks, if the sagittal suture was obliterated or completely closed, identifying the location of either bregma or lambda may be challenging, thus changing the placement of calipers for the measurement. It is possible that this measurement carries enough weight to alter an ancestry estimation, as was seen for IND12 with JHE being the only observer to disagree. Other cases of disagreement, however, did not result in a change in ancestry estimations.

Interorbital breadth (DKB) was a measurement listed by Langley et al. (2018) as having high observer error rates (relative technical error of measurement [TEM] > 2.0 which is above accepted values by Langley et al. [2018]); however, there were only three cases of disagreement in this study, all of which were from the author's second measurements. Furthermore, not all observers (i.e. ARM) had a disagreement with this measurement. This may be a result of the level of experience. However, despite agreement with this measurement, the author's discrepancies were not dissimilar to intra-observer error rates observed by Langley et al. (2018). DKB had high degrees of intra-observer error for observers in Langley et al.'s study (2018) who did not have more than ten years of experience. This may explain HBD's discrepancies, but does not explain why other observers with similar or less experience (i.e. SLP and JHE) had no instance of disagreement.

Ancestry estimations predicted by FORDISC 3.1 (Jantz and Ousley 2010a) only changed from the original estimation once out of three times where there was a disagreement in DKB measurements. The one individual where the ancestry did change (IND14) also used the MDH

measurement which was disagreed on by all observers (i.e. SLP, JHE and HBD). As such, it is possible that the change in ancestry estimation for this person is due to the changing MDH measurements.

5.3.2 Error with Mandibular Measurements

Byrnes et al. (2017) conducted a study to evaluate the observer error with both morphoscopic features on the mandible as well as their measurements. Similar to the study by Langley et al. (2018), there were four observers used, though three in Byrnes et al. (2017) held Ph.D.'s and had more experience than the fourth, who was an undergraduate. Byrnes et al. (2017) found that there was not a high degree of observer error with any of the measurements, though there were cases of disagreement for morphoscopic features (Byrnes et al. 2017). Metric features that had the lowest level of agreement included the mandibular body breadth at the mental foramen (TML) and the mandibular body breadth at the M2/M3 junction (TML23).

In the present study, there was only one case of a disagreement with TML23 (IND11) and no instances of disagreement for TML. Diagrams showing how to take each specific measurement perhaps diminished such cases of disagreement (Berg 2015:53). IND11 did experience changes in ancestry estimation from the original measurements; however, this was not unique to HBD as each observer had different estimations predicted for this individual (see above).

Despite not having difficulties for Byrnes et al. (2017), the measurement, maximum ramus height, (XRH) was the only measurement with large levels of disagreement. Given by Buikstra and Ubelaker (1994) and borrowed by Berg (2015), this measurement is the "direct distance from the highest point on the mandibular condyle to gonion" (Buikstra and Ubelaker 1994:78). Identifying the location of gonion was likely challenging as observers may have placed

the sliding calipers too far underneath the corpus of the mandible or had difficulty keeping the calipers perfectly on the landmark.

FORDISC 3.1 runs did not use XRH, as the Guatemalan male sample does not have enough recordings of this measurement to use (Jantz and Ousley 2010b:15-16). However, it is used in (hu)MANid (Berg and Kenyhercz 2017) for all composite groups, thereby possibly influencing how an unknown individual's ancestry may be estimated. For IND40.1, discrepancies in ancestry estimations between observers is possibly due to the observer taking the first measurements incorrectly and other observers demonstrated this.

As IND6 and IND40.2's XRH measurement disagreements did not contribute to an ancestry estimation changing between observers, it is possible that (hu)MANid does not weigh XRH heavily enough to generate estimations. Its impact may be more crucial for sex estimations, which this program provides.

5.5 Limitations of Study

5.5.1 Limitations Specific to Observer Error

Occasional instances of illegibility of measurements may have caused an incorrect entering of actual measurements observed into the programs. If this is the case, it could explain the changes in ancestry estimations. Incorrect measurements but legible recordings are also possible, thus also impacting how the programs interpret measurements. Incorrect readings of calipers or not fully understanding definitions of landmarks or instructions on how to take the measurement could cause incorrect measurements. Inter-observers and the intra-observer also failed in some instances to take measurements when they were available or neglected examining entire skeletal elements. Evaluations of every possible measurement and ancestry estimation were not done as a result.

5.5.2 Other Limitations

Not having a sample of positively identified individuals at the ACCO was the major source of limitation for this study as the true accuracy of any of the methods applied was impossible to test. For the one individual who is positively identified, their metric data will look to be submitted to the FDB. Though not ideal, the sample from this thesis is not unlike many that would be found in other medico-legal collections. This fact expounds the need of documenting and reporting collections similar to those stored at ACCO. Moreover, a larger documented sample is needed to continue to test the comparability of these methods.

Further, due to a lack of software and/or programming, using certain statistical analyses (e.g. k-nearest neighbor, naïve Bayesian) to estimate ancestry in this thesis was not possible. All 17 scores for all individuals are located in Appendix Table A-1 for readers to use when software/programming is available.

In some of the individuals observed, taphonomic damage prevented observation of them through multiple methods (IND16.2, IND28, IND30, IND 31, IND 32, IND36 and IND50.2). The lack of metric data prohibited the use of FORDISC 3.1 as well as (hu)MANid in several cases, and damage to morphoscopic features rendered the use of OSSA scores and other discriminant function analyses (Hefner 2015) impossible.

The inability to use the five-trait function in any of the DFA (Hefner 2015) was also challenging, resulting in the author resorting to using the three-trait function if the seven-trait function was also impossible to use. The three-trait function did have reasonable cross-validation percentages for both the three-way DFA (63.1%) and the two-way DFA (85.4%). However, using five morphoscopic traits would mean recognizing more variation which results in a higher

cross-validation percentage in one instance (three-way DFA increases to 71.0% while two-way DFA remains 85.4%).

Some individuals also were challenging to interpret (e.g. IND16.1). Discussed briefly in Chapter IV, (hu)MANid estimated this individual to be "Indeterminate" (Table 4.13). Reasons for this are unknown with secondary estimations of this individual also different between runs (Tables 4.7, 4.8, 4.9, 4.11, 4.12) where even sex estimations are not always constant. It is possible that this individual's ancestral population is not included in (hu)MANid's reference databank, thereby making a confident estimation nearly impossible.

(hu)MANid also demonstrated cases where the Euclidean Distance value and the posterior probability did not agree (see Chapter IV for counts between LDA and MDA). Though (hu)MANid uses this Euclidean Distance as opposed to the Mahalanobis' Distance, which is employed by FORDISC 3.1, reasons for why this is occurring are unknown. There are also unknown reasons for why – when there is agreement between distance values and posterior probability – agreement is much more common when employing an LDA as opposed to an MDA.

Sample sizes in the databanks, like the one used for (hu)MANid and the FDB for FORDISC 3.1 also posed challenges. There were many cases where the sample sizes of a population were not a minimum of three times larger than the number of measurements or traits inputted into the software. While all populations available for use in FORDISC 3.1 were applied to every case as per instructions in the FORDISC 3.1 Help Version 1.48 File, this was not done when running (hu)MANid. This resulted in some populations (e.g. HISPF and NEAF) regularly being omitted from consideration. The individuals from both samples, however, originated from Guatemala and Korea respectively and likely would not be present in the ACCO. Regardless,

being able to confidently and reliably use both of these groups is necessary, particularly in places where such populations are present in larger frequencies.

The use of the (hu)MANid software proved difficult at times as well, there is currently no clear tutorial of how to use the software is provided in the program. There are some recommendations available in the literature and the American Academy of Forensic Sciences Conference in February of 2019 offered a workshop on the program; however, attendance of the author was not possible at the time. (hu)MANid was run based off of recommendations available and interpreted the outputs similarly to outputs generated by FORDISC 3.1. There were still components of the results given by the software that remain unknown, however, as Euclidean Distances values frequently disagreed with the ancestries predicted by the program and the accompanying posterior probabilities. Caution for interpreting these results is warranted and care should be taken when applying any software, including FORDISC 3.1, to human remains.

Lastly, the possible presence of Native American remains in study collections is concerning in regard to the federal law, NAGPRA. While ACCO completed all necessary requirements with the local SHPO and the Idaho State BLM, there is still a potential that other coroners' or medical examiners' offices around the country may unknowingly have such groups in their offices. More research looking into such collections is warranted by other students for not only the purposes of projects, but also to become more federally compliant.

5.6 Future Directions and Concluding Remarks

This research has pointed to concerns regarding method comparability when constructing biological profiles in forensic anthropology. Much, however, is still left to be accomplished in order to make comparability and accuracy in ancestry estimations more possible, beginning with drawing attention to which populations are present in skeletal collections and which are not.

As noted in Chapter IV, there are occasional discrepancies between sex estimations within and between methods and variation in Mahalanobis' Distance/posterior probabilities between first and second predicted ancestries in FORDISC 3.1. Variation in such outputs is extremely limited when analyzing the mandible or postcranial elements. Future research has the potential to investigate the accuracy of the mandible and/or postcranial elements' ancestry estimations from FORDISC 3.1 based on these findings.

In this study, agreement with the chosen ancestry estimation methods occurred most often when the unknown individual had an ancestry estimation of American White ancestry (Table 4.18). Other ancestry estimates besides American White from ACCO are present as well (e.g. American Black, broadly Native American, broadly Asian, broadly Hispanic, etc.). Their disproportionate presence (particularly the high number of Native Americans) within the collection may result from the manner in which remains had been collected and curated in Idaho. Skeletonized individuals that fall into the possession of coroners'/medical examiners' offices are more likely to have died from non-natural manners (e.g. suicide, homicide, accidental, etc.) (Komar and Grivas 2008:228). The concern then becomes a question of why a disproportionately high number of minority groups are falling victim to such manners of death. Further discussion on such a topic may look to not only find the cause behind this phenomenon but may also look to improve the quality of life for such groups.

More research using other methods to continue testing comparability is necessary, whether with the ACCO collection or perhaps through other coroners'/medical examiners' offices. Though the methods chosen for the purposes of this study did not agree regularly, this study does succeed in broadcasting this dilemma. Likely, this is not unique to ancestry estimation methods and other methods used for the composition of the biological profile need examining. If

possible, more research regarding using k-nearest neighbor, naïve Bayesian, or canonical analysis of principal coordinates is recommended to also help assess method comparability.

Ancestry estimation methods should still be created and further developed, and using them in conjunction with one another when supplied a set of unknown remains is still recommended despite results found in this study. By using multiple methods, the forensic anthropologist should regard the patterns they generate and the limitations within them (e.g. the reference collections and databanks employed and statistical analyses used) carefully.

A continued effort to recognize human variation within reference collections is necessary for the identification of individuals in both a forensic setting as well as a bioarchaeological one. Whether through the addition of more traits available for observation (Hefner 2018) or the inclusion of other populations' metric data (Berg and Kenyhercz 2017; Jantz and Ousley 2010a), human variation is becoming more and more recognized.

For these methods to achieve greater success and become more comparable, reference collections used for their creation perhaps need to become more standardized. While there are collections that are frequently employed (e.g. William Bass Donated Collection, Terry Collection, etc.), collections with other populations aside from the traditional are not always used. Regularly employing other global populations for the creation of these methods makes it possible for these methods to become more comparable.

References Cited

Acsádi, G. and J. Nemeskéri

    1970   *History of Human Life Span and Mortality.* Akadémi Kiadó, Budapest.

Albanese, John and Saunders, Shelley R.

    2006   Is it Possible to Escape Racial Typology in Forensic Identification? In *Forensic Anthropology and Medicine*, edited by Aurore Schmitt, Eugénia Cunha, João Pinheiro, pp. 281-316. Humana Press, New York.

American Dental Association.

    2006   Tooth Eruption: The Permanent Teeth. *Journal of the American Dental Association* 137:127.

Bastir, Markus, Antonio Rosas, and P. O'Higgins

    2006   Craniofacial Levels and the Morphological Maturation of the Human Skull. *Journal of Anatomy* 209:637-654.

Berg, Gregory E.

    2015   Biological Affinity and Sex from the Mandible Utilizing Multiple World Populations. In *Biological Affinity in Forensic Identification of Human Skeletal Remains: Beyond Black and White*, edited by Gregory E. Berg and Sabrina C. Ta'ala, pp. 43-81. Taylor & Francis Group, Boca Raton, Florida.

Berg, Gregory E. and Michael W. Kenyhercz

    2017   Introducing Human Mandible Identification [(hu)MANid]: A Free, Web-Based GUI to Classify Human Mandibles. *Journal of Forensic Sciences* 62(6):1592-1598.

Betti, Lia, François Balloux, Tsunehiko Hanihara, Andrea Manica

    2010   The Relative Role of Drift and Selection in Shaping the Human Skull. *American Journal of Physical Anthropology* 141:76-82.

Blau, Soren and Douglas H. Ubelaker

    2016   *Handbook of Forensic Anthropology and Archaeology*. Routledge, New York.

Boas, Franz

    1912   Changes in the Bodily Form of Descendants of Immigrants. *American Anthropologist* 5(14):530-562.

    1931   Race and Progress. *Science* 74(1905):1-8.

Bogin, Barry

    1999   *Patterns of Human Growth*. Cambridge University Press, United Kingdom.

Boyd, Clifford and Donna C. Boyd

    2011   Theory and the Scientific Basis for Forensic Anthropology. *Journal of Forensic Sciences* 56(6):1407-1415.

Brooks, S. and J.M. Suchey

    1990   Skeletal Age Determination Based on the Os Pubis: A Comparison of the Acsadi-Nemeskeri and Suchey-Brooks Method. *Human Evolution* 5:227-238.

Buikstra, Jane E., and Douglas H. Ubelaker

    1994   *Standards: For Data Collection from Human Skeletal Remains*. Arkansas Archeological Survey Research Series: Arkansas Archeological Survey, Fayetteville.

Bulatao, Rodolfo A. and Norman B. Anderson

> 2004   Genetic Influences. In *Understanding Racial and Ethnic Differences in Health in Late Life: A Research Agenda*, edited by R.A. Bulatao and N.B. Anderson, pp. 46-53. The National Academic Press, Washington D.C.

Byrnes, Jennifer F., Michael W. Kenyhercz, Gregory E. Berg

> 2017   Examining Interobserver Reliability of Metric and Morphoscopic Characteristics of the Mandible. *Journal of Forensic Sciences* 62(4):981-985.

Carpenter, James Culver

> 1976   A Comparative Study of Metric and Non-Metric Traits in a Series of Modern Crania. *American Journal of Physical Anthropology* 45:337-344.

Caspari, Rachel

> 2003   From Types to Populations: A Century of Race, Physical Anthropology, and the American Anthropological Association. *American Anthropologist* 105(1):65-76.

Cheverud, James M., Jane E. Buikstra, Elizabeth Twichell

> 1979   Relationships between Non-Metric Skeletal Traits and Cranial Size and Shape. *American Journal of Physical Anthropology* 50:191-198.

Christensen, Angi M., Nicholas V. Passalacqua, and Eric J. Bartelink

> 2014   *Forensic Anthropology: Current Methods and Practice.* Elsevier, San Diego, California.

Corruccini, Robert S.

> 1974   An Examination of the Meaning of Cranial Discrete Traits for Human Skeletal Biological Studies. *American Journal of Physical Anthropology* 40:425-446.

DeWitte, Sharon N. and Christopher M. Stojanowski

>2015   The Osteological Paradox 20 Years Later: Past Perspectives, Future Directions. *Journal of Archaeological Research* 23:397-450.

Elliott, Marina and Mark Collard

>2009   FORDISC and the Determination of Ancestry from Cranial Measurements. *Biology Letters* 5:849-852.

Finnegan M. and S.A. McGuire

>1979   Classification Systems for Discrete Variables used in Forensic Anthropology. *American Journal of Physical Anthropology* 51:547-553.

Garn, Stanley M. and Eugene Giles

>1995   *Earnest Albert Hooton: A Biographical Memoir*. Biographical Memoirs No. 66:167-179. The National Academies Press. Washington D.C.

Gould, Stephen J.

>1980   *The Mismeasure of Man*. W.W. Norton & Company, New York, New York.

>1982   The Geometer of Race. *Discover: Science, Technology, and the Future* 1-8.

Hefner, Joseph T.

>2009   Cranial Nonmetric Variation in Estimating Ancestry. *Journal of Forensic Sciences* 54(5):985-995.

>2015   Cranial Morphoscopic Traits and the Assessment of American Black, American White, and Hispanic Ancestry. In *Biological Affinity in Forensic Identification of Human Skeletal Remains: Beyond Black and White*, edited by Gregory E. Berg and Sabrina C. Ta'ala, pp. 27-41. Taylor & Francis Group, Boca Raton, Florida.

2018   The Macromorphoscopic Databank. *American Journal of Physical Anthropology* 166(4):994-1004.

Hefner, Joseph T. and Kandus C. Linde

2018   Atlas of Human Cranial Macromorphoscopic Traits. Elsevier, San Diego, California.

Hefner, Joseph T. and Stephen D. Ousley

2006   Morphoscopic Traits and the Statistical Determination of Ancestry II. Paper presented at the 58th Annual Meeting of the American Academy of Forensic Sciences, Colorado Springs, Colorado.

2014   Statistical Classification Methods for Estimating Ancestry Using Morphoscopic Traits. *Journal of Forensic Sciences* 59(4):883-890.

Hooton, Earnest A.

1939   The American Criminal: An Anthropological Study. Harvard University Press, Cambridge, Massachusetts.

Howells, W.W.

1973   *Cranial Variation in Man: A Study of Multivariate Analysis of Patterns of Difference among Recent Human Populations*. Harvard University Press, Cambridge, Massachusetts.

1989. Skull Shapes and the Map: Craniometric Analysis in the Dispersion of Modern Homo. Harvard University Press, Cambridge, Massachusetts.

Isçan, M.Y., S.R. Loth, R.K. Wright

1984   Age Estimation from the Rib by Phases Analysis: White Males. *Journal of Forensic Science* 29:1094-1104.

1985   Age Estimation from the Rib by Phases Analysis: White Females. *Journal of Forensic Science* 30:853-863.

Jantz, Richard L. and Douglas W. Owsley

2001   Variation among Early North American Crania. *American Journal of Physical Anthropology* 114:146-155.

Jantz, Richard L., Stephen D. Ousley [does this look right for SAA?]

2010a FORDISC 3.1: Computerized Forensic Discriminant Functions. Version 3.1. The University of Tennessee, Knoxville.

2010b FORDISC Help Version 1.48 File. The University of Tennessee, Knoxville.

Johnson, Sylvester

2015   Monstrosity, Colonialism, and the Racial State. *The Journal of Nineteenth-Century Americanists* 3(1):173-181.

Klales, Alexandra R. and Michael W. Kenyhercz.

2014   Morphological Assessment of Ancestry using Cranial Macromorphoscopics. *Journal of Forensic Sciences* 60(1):13-20.

Kenyhercz, Michael W. and Gregory E. Berg

2018   Evaluating Mixture Discriminant Analysis to Classify Human Mandibles with (hu)MANid, a Free, R-Based GUI. In *New Perspectives in Forensic Human Skeletal Identification*, edited by Krista E. Latham, Eric J. Bartelink and Michael Finnegan, pp. 35-43. Elsevier, San Diego, California.

Komar, Debra A. and Christopher Grivas

2008   Manufactured Populations: What do Contemporary Reference Skeletal

Collections Represent? A Comparative Study Using the Maxwell Museum Documented

Collection. *American Journal of Physical Anthropology* 137:224-233.

Langley, Natalie R., Lee Meadows Jantz, Shauna McNulty, Heli Maijanen, Stephen D. Ousley,

Richard L. Jantz

2018   Error Quantification of Osteometric Data in Forensic Anthropology. *Forensic*

Science International 287:183-189.

Langley, Natalie R., Lee Meadows Jantz, Stephen D. Ousley, Richard L. Jantz, George Milner

2016   Data Collection Procedures for Forensic Skeletal Material 2.0. University of

Tennessee and Lincoln Memorial University, Knoxville and Harrogate, Tennessee.

Lewontin, Richard C.

1972   The Apportionment of Human Diversity. *Evolutionary Biology* 6:381-398.

2006. Confusion about Human Races. Electronic document,

http://raceandgenomics.ssrc.org/Lewontin/, accessed October 23, 2017.

Lovejoy, C.O., R.S. Meindel, R. Mensforth, T.J. Barton

1985   Chronological Metamorphosis of the Auricular Surface of the Ilium: A New

Method for the Determination of Adult Skeletal Age at Death. *American Journal of*

Physical Anthropology 68:15-28.

Messer, Diana B.

2013   Human Craniometric Sexual Dimorphism: Assessing Rensch's Rule and Its

Implications for Forensic Classification. Master's thesis, Applied Forensic Sciences,

Mercyhurst University, Erie, PA.

Messer, Diana B. and Sara M. Getz

>   In Press   The Effect of Sex Misclassification on the Skeletal Biological Profile. In
>
>   *Sex Estimation of the Human Skeleton: History, Methods, and Emerging Techniques*,
>
>   edited by A. Klales.

Milner, G.R.

>   1992   Determination of Skeletal Age and Sex: A Manual Prepared for the Dickson
>
>   Mounds Reburial Team. Ms. on file, Dickson Mounds Museum, Lewiston, Illinois.

Molnar, Stephen

>   2006   *Human Variation: Races, Types, and Ethnic Groups*. Pearson Education, New
>
>   Jersey.

Moore-Jansen, P.M., Stephen D. Ousley, Richard L. Jantz

>   1994   *Data Collection Procedures for Forensic Skeletal Material*. Report of
>
>   Investigations 48. University of Tennessee, Knoxville.

Morton, Samuel G.

>   1839   *Crania Americana, or A Comparative View of the Skulls of Various Aboriginal*
>
>   *Nations of North and South America: to which is Prefixed an Essay on the Varieties of*
>
>   *the Human Species*. J. Dobson, Philadelphia, Pennsylvania.

NAGPRA

>   1990 Native American Graves Protection and Repatriation Act.

Nakhaeizadeh, Sherry, Itiel E. Dror, Ruth M. Morgan

>   2013   Cognitive Bias in Forensic Anthropology: Visual Assessment of Skeletal Remains
>
>   is Susceptible to Confirmation Bias. *Science and Justice* 54:2018-214.

National Conference of State Legislatures

    2018   List of Federal and State Recognized Tribes,

            www.ncsl.org/research/state-tribal-institute/list-of-federal-and-state-recognized-

            tribes.aspx#Id, accessed October 2018.

Ousley, Stephen, Richard L. Jantz, and Joseph T. Hefner

    2018   From Blumenbach to Howells: The Slow, Painful Emergence of Theory through

            Forensic Race Estimation. In *Forensic Anthropology: Theoretical Framework and*

            Scientific Basis, edited by C. Clifford Boyd Jr. and Donna C. Boyd, pp. 67-97. Wiley,

            Hoboken, New Jersey.

Plemons, Amber and Joseph T. Hefner

    2016   Ancestry Estimation Using Macromorphoscopic Traits. *Academic Forensic*

            Pathology Journal 6(3):400-412.

Phenice, T.

    1969   A Newly Developed Visual Method of Sexing in the Os Pubis. *American Journal*

            of Physical Anthropology 8:679-684.

Rösing, F.W., M. Graw, B. Marré, S. Ritz-Timme, M.A. Rothschild, K. Rötzscher, A.

Schmeling, I. Schröder, G. Geserick.

    2007   Recommendations for the Forensic Diagnosis of Sex and Age from Skeletons.

            Journal of Comparative Human Biology 58:75-89.

Richtsmeier, Joan T., James M. Cheverud and Jane E. Buikstra

    1984   The Relationship Between Cranial Metric and Nonmetric Traits in the Rhesus

            Macaques from Cayo Santiago. *American Journal of Physical Anthropology* 64:213-222.

Rushton, J. Philippe

    1997   Race, Intelligence, and the Brain: The Errors and Omissions of the 'Revised'

    Edition of S.J. Gould's *The Mismeasure of Man. Personality and Individual Differences*

    23(1):169-180.

Sauer, Norman J.

    1992   Forensic Anthropology and the Concept of Race: If Races Don't Exist, Why are

    Forensic Anthropologists so Good at Identifying Them? *Social Science & Medicine*

    34(2)107-111.

Sauer, Norman J., Jane C. Wankmiller, and Joseph T. Hefner

    2016   Assessment of Ancestry and the Concept of Race. In *Handboook of Forensic*

    Anthropology, 2nd Edition, edited by Soren Blau and Douglas H. Ubelaker, pp. 243-260.

    Routledge, New York City, New York.

Smedley, Audrey; and Brian D. Smedley

    2005   Race as Biology is Fiction, Racism as a Social Construct is Real. *American*

    Psychologist 60(1):16-26.

Spradley, Katherine M., Jantz, Richard L., Robinson, Alan; Peccerelli, Fredy

    2008   Demographic Change and Forensic Identification: Problems in Metric

    Identification of Hispanic Skeletons. *Journal of Forensic Sciences* 53(1):21-28.

Stephens, Christianne

    2000   A Measure of Truth? An Examination of the History, Methodology, Theory and

    Practice of Craniometric Analysis, with Special Focus on its Application in Forensic

    Identification. Totem: *The University of Western Ontario Journal of Anthropology*

    8(1):84-108.

Tasa, Guy and Juliette Vogel

>2016   The Use of Howells' Dataset in Determining Ethnicity in Pacific Northwest
>Crania: Implications for Kennewick Man. Paper presented at the 69th Annual Northwest
>Anthropological Conference, Tacoma, Washington.

Tishkoff, Sarah A.; and Kenneth K. Kidd

>2004   Implications of Biogeography of Human Populations for 'Race' and Medicine.
>*Nature Genetics* 36(11):521-527.

Ubelaker, Douglas H.

>1989a *Human Skeletal Remains.* 2nd Edition. Taraxacum Press, Washington D.C.

>1989b The Estimation of Age at Death from Immature Human Bone. In *Age Markers in
>the Human Skeleton*, edited by M.Y. Isçan, pp. 55-70. Charles C. Thomas, Springfield,
>Illinois.

University of Tennessee Knoxville

>Forensic Anthropology Center. Forensic Anthropology Data Bank (FDB)
>   http://fac.utk.edu/background, accessed October 2018.

US Census Bureau

>2017   American Community Survey, Table B03002; Idaho Population by Race
>http://worldpopulationreview.com/states/idaho-population/#byRace, accessed
>2019.

>2017   American Community Survey, Table S0101; Idaho Population by Age
>http://worldpopulationreview.com/states/idaho-population/#byAge, accessed
>2019.

Wagner, Jennifer K.; Joon-Ho Yu, Jayne O. Ifekwunigwe, Tanya M. Harrell, Michael Bamshad,
Charmaine D. Royal

  2016   Anthropologists' Views on Race, Ancestry, and Genetics. *American Journal of Physical Anthropology* 162:318-327.

Wheat, Amber B.

  2009   Assessing Ancestry through Nonmetric Traits of the Skull: A Test of Education and Experience. Master's Thesis, Department of Anthropology, Texas State University-San Marcos, San Marcos.

Williams, Frank L., Robert L. Belcher, George J. Armelagos

  2005   Forensic Misclassification of Ancient Nubian Crania: Implications for Assumptions about Human Variation. *Current Anthropology* 46(2):340-346.

Wilson, Rebecca J., Nicholas P. Herrmann, Lee Meadows Jantz

  2010   Evaluation of Stature Estimation from the Database for Forensic Anthropology. *Journal of Forensic Sciences* 55(3):684-689.

Wolf, Eric

  1992   Perilous Ideas: Race, Culture, People. *Current Anthropology* 35(1):1-12.

Wood, James W., Milner, George R., Harpending, Henry C., Weiss, Kenneth M.

  1992   The Osteological Paradox: Problems of Inferring Prehistoric Health from Skeletal Samples. *Current Anthropology* 33(4):343-370.

Abbreviations for Hefner (2018) (Table A-1)

ANS: Anterior Nasal Aperture
INA: Inferior Nasal Aperture
IOB: Interorbital Breadth
MT: Malar Tubercle
NAS: Nasal Aperture Shape
NAW: Nasal Aperture Width
NBC: Nasal Bone Contour
NBS: Nasal Bone Shape
NO: Nasal Overgrowth
NFS: Nasofrontal Suture
OBS: Orbital Shape
PBD: Post-bregmatic Depression
PZT: Posterior Zygomatic Tubercle
SPS: Supranasal Suture
TPS: Transverse Palatine Suture
PS: Palate Shape
ZS: Zygomaticomaxillary Suture Course

Table A-1. Individual Macromorphoscopic Scores using Hefner (2018).

Case Number	ANS	INA	IOB	MT	NAS	NAW	NBC	NBS	NO	NFS	OBS	PBD	PZT	SPS	TPS	PS	ZS
IND1	3	5	1	0	3	1	3	N/A	1	1	1	1	2	0	4	1	1
IND2	2	2	1	1	1	1	-	N/A	-	3	1	1	2	2	1	1	2
IND3	2	1	2	-	3	3	4	N/A	-	1	1	0	3	2	1	1	2
IND4	-	2	1	1	3	2	0	N/A	-	1	1	0	1	0	2	4	1
IND5	-	-	-	-	-	-	-	N/A	-	2	-	0	-	1	-	-	-
IND6	2	5	1	1	2	1	3	N/A	1	1	3	0	1	0	1	3	1
IND7	-	2	1	0	3	2	1	N/A	-	4	1	0	1	0	2	1	0
IND8	2	3	2	1	2	3	1	N/A	1	2	1	0	1	0	3	2	0
IND9	-	-	3	-	-	1	3	N/A	-	4	2	0	-	0	-	-	-
IND10	3	3	3	1	2	2	1	N/A	1	1	3	0	1	0	3	3	1
IND11	2	4	2	3	3	2	3	N/A	1	2	1	0	2	2	2	3	0
IND12	-	-	2	-	-	-	-	N/A	-	-	-	0	-	0	-	-	-
IND13	-	3	2	3	2	2	4	N/A	-	2	2	0	1	0	2	-	0
IND14	2	3	2	0	2	1	3	N/A	1	1	1	0	1	0	1	3	2
IND15	-	-	2	0	3	2	0	N/A	-	2	1	0	1	0	-	4	2
IND16.1	-	-	-	-	-	-	-	-	-	-	-	-	3	-	-	-	-
IDN23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
IND28	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
IND29	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
IND30	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
IND34	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
IND35	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
IND36	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
IND37	-	3	3	2	2	2	1	N/A	-	2	1	0	3	1	1	1	1
IND38	-	2	1	1	1	1	4	N/A	-	1	2	0	2	2	1	2	0
IND39	-	-	2	-	-	-	-	N/A	-	2	-	-	0	0	-	-	-
IND45	3	4	3	0	2	2	3	N/A	1	1	1	1	2	2	1	3	0
IND46	3	3	3	1	2	2	1	N/A	1	1	3	0	1	2	3	-	0
IND47	2	3	2	2	3	2	2	N/A	-	4	2	0	1	0	-	-	0
IND50.1	2	3	1	2	3	3	3	N/A	-	4	1	-	1	2	1	-	1
IND51	2	2	1	3	3	2	-	N/A	-	-	2	-	1	0	2	2	0
IND52	-	1	1	0	2	2	0	N/A	-	4	3	1	2	0	1	4	0
IND53	2	3	3	3	3	3	0	N/A	1	3	3	0	3	2	2	-	0
IND54	3	4	1	0	2	1	3	N/A	0	-	3	-	0	2	3	4	0
IND55	-	5	3	0	3	-	3	N/A	-	2	1	0	0	0	-	1	0
IND56	3	4	1	0	3	1	4	N/A	-	-	3	1	2	0	3	3	1
IND57	2	2	2	0	2	1	4	N/A	-	1	2	0	1	2	1	2	2

N/A: Specific tool required that was not available for use

Abbreviations for Berg (2015) Morphoscopics (Table A-2)

CS: Chin Shape
LBM: Lower Border of Mandible
ARS: Ascending Ramus Shape
GAF: Gonial Angle Flare
MT: Mandibular Torus
PREI: Posterior Ramus Edge Inversion

Table A-2. Individual Morphoscopic Scores using Berg (2015)

Case Number	CS	LBM	ARS	GAF	MT	PREI
IND6	4	1	2	5	2	2 or 3
IND8	1	1	1	4	2	3
IND10	4	2	1	3	2	2
IND11	1	2	2	2	2	1
IND16.1	4	2	2	4	1	4
IND23	1	1	1	2	1	1
IND25	2	1	2	2	1	2
IND26	1	2	1	3	1	3
IND27	3	1	2	2	1	3
IND28	1	2	-	-	1	-
IND29	4	2	-	4	1	2
IND30	3	1	-	-	2	-
IND31	3	3	-	2	1	2
IND32	3	3	-	5	2	4
IND33	1	2	-	5	1	1
IND34	-	1	1	3	2	4
IND35	2	2	-	3	1	4
IND36	3	2	-	5	2	1
IND37	1	1	2	4	2	1
IND40.1	1	3	2	3	1	1
IND40.2	1	1	1	2	1	4
IND47	1	3	1	3	1	2
IND48	3	1	1	4	2	2
IND49	1	3	2	4	1	1
IND50.2	4	2	1	1	1	2
IND54	1	2	1	2	1	1
IND55	4	1	2	5	2	2
IND56	3	2	1	3	2	1

Index for Cranial Measurements (Table A-3)
1. Maximum Cranial Length
2. Maximum Cranial Breadth
3. Bizygomatic Breadth
4. Basion-Bregma Height
5. Cranial Base Length
6. Basion-Prosthion Length
7. Maxillo-Alveolar Breadth
8. Maxillo-Alveolar Length
9. Biauricular Breadth
10. Nasion-Prosthion Height
11. Minimum Frontal Breadth
12. Upper Facial Height
13. Nasal Height
14. Nasal Breadth
15. Orbital Breadth (Left and Right)
16. Orbital Height (Left and Right)
17. Biorbital Breadth
18. Interorbital Breadth
19. Frontal Chord
20. Parietal Chord
21. Occipital Chord
22. Foramen Magnum Length
23. Foramen Magnum Breadth
24. Mastoid Height (Left and Right)

Table A-3. Cranial Measurements of Sample in mm.

Case Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15 L	15 R	16 L	16 R	17	18	19	20	21	22	23	24 L	24 R
IND1	183	136	127	143	101	-	-	-	120	-	95	105	53	24	38	36	35	33	88	(22)	116	123	96	39	34	34	38
IND2	178	130	-	127	97	(95)	60	54	120	72	89	103	49	22	35	36	33	32	93	21	108	111	97	34	28	29	29
IND3	183	141	145	129	105	(99)	67	(52)	135	72	94	108	59	30	39	39	37	36	94	25	105	108	94	37	31	36	36
IND4	180	137	-	124	93	98	56	49	126	59	93	103	43	24	38	38	31	31	96	22	107	113	90	32	28	25	24
IND5	183	136	-	-	-	-	-	-	(116)	-	96	-	-	-	-	-	-	-	-	-	(107)	(116)	-	-	-	-	-
IND6	195	138	135	153	104	95	65	57	130	83	101	112	57	23	44	43	38	37	104	19	128	122	105	40	35	36	38
IND7	-	119	-	-	-	-	61	(52)	-	(63)	88	100	-	25	(35)	(36)	31	33	94	-	105	109	-	-	-	-	26
IND8	172	136	122	127	94	91	-	53	120	64	94	102	50	27	38	38	32	32	94	23	100	112	96	34	30	24	27
IND9	180	131	-	128	95	-	-	-	(103)	-	88	94	-	-	35	35	36	37	88	20	109	112	-	-	-	-	-
IND10	(181)	(135)	126	139	111	109	66	55	123	74	-	106	53	23	(38)	(39)	33	34	97	-	(110)	(110)	95	35	34	31	31
IND11	185	133	138	(147)	100	(95)	66	(49)	131	(74)	97	108	59	24	40	40	36	36	101	23	113	112	104	37	31	28	27
IND12	187	139	-	-	-	-	-	-	-	-	95	(106)	-	-	-	-	-	-	-	-	(124)	116	-	-	-	-	30
IND13	189	140	-	137	-	-	-	-	-	-	100	114	-	27	-	-	-	36	-	-	117	106	100	37	30	-	(32)
IND14	184	147	127	132	98	-	65	-	121	-	100	103	56	23	39	37	35	33	94	(24)	113	117	104	34	31	34	32
IND15	181	133	-	129	101	-	-	-	127	-	91	102	-	-	37	-	33	-	-	-	114	100	100	32	27	31	34
IND16.1	-	-	-	-	-	-	64	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	30	23
IND23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
IND28	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
IND29	-	-	-	-	-	-	65	(48)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
IND30	-	-	-	-	-	-	65	59	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
IND34	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
IND35	-	-	-	-	-	-	63	56	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
IND36	-	-	-	-	-	-	62	53	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
IND37	187	141	154	125	100	(101)	-	-	138	(74)	100	117	(53)	28	43	43	38	38	109	27	110	(104)	95	33	26	34	30
IND38	183	129	-	130	97	-	63	(47)	122	-	93	102	50	24	35	37	35	36	90	21	113	115	87	39	33	-	28
IND39	(174)	144	-	123	-	-	-	-	122	-	92	(97)	-	-	-	-	-	-	-	-	-	109	87	32	30	18	23

Case Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15		16		17	18	19	20	21	22	23	24	
															L	R	L	R								L	R
IND45	184	134	124	132	105	104	(61)	52	115	67	100	105	52	23	36	38	32	31	95	29	104	126	97	35	29	29	30
IND46	187	142	137	138	101	(92)	-	(49)	132	(73)	106	114	53	23	41	41	35	36	102	25	116	(114)	(99)	39	33	32	34
IND47	170	125	-	-	-	-	-	-	113	67	-	-	47	26	-	37	-	34	-	23	100	107	91	-	-	26	25
IND50.1	182	-	-	-	100	(98)	-	-	-	(61)	-	-	(47)	-	38	-	33	-	(19)	-	-	102	30	27	27	27	-
IND51	-	-	-	-	-	-	58	48	-	(67)	90	94	(48)	22	34	35	34	34	89	23	(100)	-	-	-	-	-	-
IND52	179	138	129	123	98	(101)	-	52	121	(68)	92	106	-	23	40	40	34	34	96	22	105	(107)	98	(28)	23	27	27
IND53	190	136	(144)	131	105	-	-	-	138	-	96	114	(53)	27	39	40	36	37	104	27	110	(106)	100	(34)	27	29	30
IND54	204	148	-	147	116	-	-	-	128	-	107	113	62	25	43	43	35	35	101	22	125	111	119	39	32	35	36
IND55	(180)	125	-	128	103	-	-	-	115	-	93	(104)	(50)	-	38	-	30	-	-	-	109	(112)	96	32	29	30	31
IND56	191	(151)	-	136	103	99	-	96	-	80	101	106	60	22	41	39	38	37	97	22	115	99	108	35	31	28	30
IND57	153	153	-	129	95	(92)	-	(51)	127	68	93	(104)	51	24	-	36	37	37	-	24	105	93	88	33	26	(29)	30

()'s: Aspect of measurement was damaged and measurement was used with caution

Measurements (in mm) for Postcrania specific to Individual.

Table A-4. Postcranial Measurements for IND6 in mm.

IND6	Left	Right	Single Element
Scapula			
Height	160	162	
Breadth	107	108	
Humerus			
Maximum Length	336	340	
Epicondylar Breadth	68	68	
Max. Vertical Head Diameter	50	50	
Max. Midshaft Diameter	23	24	
Min. Midshaft Diameter	19	17	
Radius			
Maximum Length	262	258	
Maximum/Transverse Midshaft Diameter	15	15	
Minimum/Anterior-Posterior Midshaft Diameter	12	13	
Ulna			
Maximum Length	(279)	277	
Maximum/Dorso-Volar Midshaft Diameter	18	18	
Minimum/Transverse Midshaft Diameter	13	13	
Physiological Length	246	241	
Minimum Circumference	-	-	
Sacrum			
Anterior Height			112
Anterior Breadth			116
Transverse Diameter S1			(60)
Anterior-Posterior Diameter S-1			38
Innominate			
Maximum Height	220	218	
Maximum Iliac Breadth	167	166	
Femur			
Maximum Length	-	490	
Bicondylar Length	-	483	
Epicondylar Breadth	-	89	
Maximum Head Diameter	-	44	
Transverse Subtrochanteric Diameter	-	34	
Anterior-Posterior Subtrochanteric Diameter	-	29	
Anterior-Posterior Midshaft Diameter	-	29	
Transverse Midshaft Diameter	-	30	
Midshaft Circumference	-	98	

IND6	Left	Right	Single Element
Tibia			
Condylo-Malleolar Length/Maximum. Length	412	412	
Maximum Proximal Epicondylar Breadth	81	81	
Distal Epicondylar Breadth	55	52	
Maximum Midshaft Diameter at Nutrient Foramen	36	38	
Minimum Midshaft Diameter at Nutrient Foramen	24	24	
Midshaft Circumference at Nutrient Foramen	100	103	
Fibula			
Maximum Length	410	-	
Maximum Midshaft Diameter	14	-	
Calcaneus			
Maximum Length	-	-	
Middle Breadth	43	45	

()'s: Aspect of measurement was damaged and measurement was used with caution

Table A-5. Postcranial Measurements for IND16.1 in mm.

IND16.1	L	R	Single Element
Scapula			
Height	-	-	
Breadth	-	104	
Humerus			
Maximum Length	321	-	
Epicondylar Breadth	64	63	
Maximum Vertical Head Diameter	46	-	
Maximum Midshaft Diameter	21	(21)	
Minimum Midshaft Diameter	15	(16)	
Sacrum			
Anterior Height			-
Anterior Breadth			116
Transverse Diameter S-1			60
Anterior-Posterior Diameter S-1			-
Innominate			
Maximum Length	218	-	
Maximum Iliac Breadth	163	159	
Femur			
Maximum Length	460	-	
Bicondylar Length	459	-	
Epicondylar Breadth	81	80	
Maximum Head Diameter	47	46	
Transverse Subtrochanteric Diameter	29	31	
Anterior-Posterior Subtrochanteric Diameter	25	25	
Transverse Midshaft Diameter	28	(28)	
Anterior-Posterior Midshaft Diameter	28	(25)	
Midshaft Circumference	90	91	
Tibia			
Condylo-Malleolar Length/Maximum Length	376	376	
Maximum Proximal Epicondylar Breadth	77	76	
Distal Epicondylar Breadth	55	54	
Maximum Midshaft Diameter at Nutrient Foramen	36	35	
Minimum Midshaft Diameter at Nutrient Foramen	18	21	
Midshaft Circumference	95	94	
Fibula			
Maximum Length	-	362	
Maximum Midshaft Diameter	16	17	

()'s: Aspect of measurement was damaged and measurement was used with caution

Table A-6. Postcranial Measurements for IND16.2 in mm.

IND16.2	L	R
Radius		
Maximum Length	242	-
Maximum/Transverse Midshaft Diameter	-	-
Minimum/Anterior-Posterior Midshaft Diameter	-	-
Ulna		
Maximum Length	-	-
Maximum/Dorso-Volar Midshaft Diameter	-	(14)
Minimum/Transverse Midshaft Diameter	-	(12)
Physiological Length	233	-
Minimum Circumference	-	-

()'s: Aspect of measurement was damaged and measurement was used with caution

Table A-7. Postcranial Measurements for IND23 in mm

IND23	L	R
Scapula		
Height	142	-
Breadth	-	-
Humerus		
Maximum Length	-	-
Epicondylar Breadth	-	61
Maximum Vertical Head Diameter	-	-
Maximum Midshaft Diameter	-	-
Minimum Midshaft Diameter	-	-
Innominate		
Maximum Height	-	-
Maximum Breadth	148	-
Femur		
Maximum Length	-	-
Bicondylar Length	-	-
Epicondylar Breadth	-	-
Maximum Head Diameter	-	(47)
Transverse Subtrochanteric Diameter	29	33
Anterior-Posterior Subtrochanteric Diameter	25	25
Transverse Midshaft Diameter	-	-
Anterior-Posterior Midshaft Diameter	-	-
Midshaft Circumference	-	-
Tibia		
Condylo-Malleolar Length/Maximum Length	-	-
Maximum Proximal Epicondylar Breadth	-	-
Distal Epicondylar Breadth	48	(50)
Maximum Midshaft Diameter at Nutrient Foramen	-	37
Minimum Midshaft Diameter at Nutrient Foramen	-	22
Midshaft Circumference	-	95
Calcaneus		
Maximum Length	-	-
Middle Breadth	-	43

()'s: Aspect of measurement was damaged and measurement was used with caution

Table A-8. Postcranial Measurements for IND47 in mm.

IND47	L	R
Humerus		
Maxium Length	-	-
Epicondylar Breadth	-	-
Maximum Vertical Head Diameter	-	(37)
Maximum Midshaft Diameter	-	-
Minimum Midshaft Diameter	-	-
Radius		
Maximum Length	(225)	-
Maximum/Transverse Midshaft Diameter	12	-
Minimum/Anterior-Posterior Midshaft Diameter	10	-
Ulna		
Maximum Length	-	(248)
Maximum/Dorso-Volar Midshaft Diameter	-	13
Minimum/Transverse Midshaft Diameter	-	11
Physiological Length	-	(219)
Minimum Circumference	-	-
Tibia		
Condylo-Malleolar Length/Maximum Length	-	-
Maximum Proximal Epicondylar Breadth	-	-
Distal Epicondylar Breadth	-	-
Maximum Midshaft Diameter at Nutrient Foramen	28	-
Minimum Midshaft Diameter at Nutrient Foramen	17	-
Midshaft Circumference	78	-

()'s: Aspect of measurement was damaged and measurement was used with caution

Table A-9. Postcranial Measurements for IND54 in mm. (Not Run through FORDISC 3.1)

IND54	L	R
Calcaneus		
Maximum Length	84	-
Middle Breadth	41	-

Table A-10. Postcranial Measurements for IND55 in mm.

IND55	L	R	Single Element
Clavicle			
Maximum Length	154	156	
Maximum/Anterior-Posterior Midshaft Diameter	16	16	
Minimum/Superior-Inferior Midshaft Diameter	11	11	
Scapula			
Height	-	(175)	
Breadth	-	97	
Humerus			
Maximum Length	324	-	
Epicondylar Breadth	58	-	
Maximum Vertical Head Diameter	49	-	
Maximum Midshaft Diameter	26	-	
Minimum Midshaft Diameter	21	-	
Radius			
Maximum Length	236	(236)	
Maximum/Transverse Midshaft Diameter	19	17	
Minimum/Anterior-Posterior Midshaft Diameter	13	14	
Ulna			
Maximum Length	260	(258)	
Maximum/Anterior-Posterior Midshaft Diameter	21	20	
Minimum/Transverse Midshaft Diameter	14	16	
Physiological Length	220	(222)	
Minimum Circumference	-	-	
Sacrum			
Anterior Height			111
Anterior Breadth			119
Transverse Diameter S-1			56
Anterior-Posterior Diameter S-1			
Innominate			
Maximum Height	232	226	
Maximum Iliac Breadth	169	172	
Femur			
Maximum Length	432	448	
Bicondylar Length	432	445	
Epicondylar Breadth	(83)	77	
Maximum Head Diameter	50	53	
Transverse Subtrochanteric Diameter	32	28	
Anterior-Posterior Subtrochanteric Diameter	30	30	
Transverse Midshaft Diameter	31	29	
Anterior-Posterior Midshaft Diameter	31	29	

IND55	L	R	Single Element
Midshaft Circumference	100	95	
Tibia			
Condylo-Malleolar Length/Maximum Length	349	350	
Maximum Proximal Epicondylar Breadth	77	(72)	
Distal Epicondylar Breadth	54	(51)	
Maximum Midshaft Diameter at Nutrient Foramen	34	(33)	
Minimum Midshaft Diameter at Nutrient Foramen	26	(25)	
Midshaft Circumference	97	(95)	
Fibula			
Maximum Length	344	-	
Maximum Midshaft Diameter	17	-	
Calcaneus			
Maximum Length	82	-	
Middle Breadth	46	-	

()'s: Aspect of measurement was damaged and measurement was used with caution

Appendix Table B-1. (hu)MANid (Berg and Kenyhercz 2017) Outputs using Measurements Taken by SLP and Linear Discriminant Analysis (LDA).

Case Number	# of Meas. Used	# of Pops. Used	Predicted Ancestry	Posterior Probability	Second Predicted Ancestry	Posterior Probability	Distances to Centroid
IND6	9	4	WM	0.952	SEAM	0.035	SEAM(1.658) --> 1.902
IND27	9	9	HISPM	0.705	BM	0.066	0.480 --> SEAF(0.784)
IND32	3	12	HISPF	0.385	BF	0.221	0.919 --> AIF(1.074)
IND33	5	10	NEAM	0.257	HISPM	0.232	0.392 --> WM(0.962)
IND40.1	8	3	AIF	0.546	BF	0.267	0.726 --> 1.400
IND40.2	8	3	AIM	0.538	AIF	0.448	2.876 --> 2.939

Appendix Table B-2. (hu)MANid (Berg and Kenyhercz 2017) Outputs using Measurements Taken by SLP and Mixture Discriminant Analysis (MDA).

Case Number	# of Meas. Used	# of Pops. Used	Predicted Ancestry	Posterior Probability	Second Predicted Ancestry	Posterior Probability	Distances to Centroid
IND6	9	4	WM	0.872	SEAM	0.124	2.199 --> 2.600
IND27	9	9	HISPM	0.542	SEAM	0.128	AIF(0.663)--->1.149
IND32	3	12	BF	0.211	HISPF	0.176	HISPF(1.450)--->1.837
IND33	5	10	HISPM	0.371	SEAM	0.186	NEAM(1.299)--->2.366
IND40.1	8	3	SEAF	0.472	AIF	0.331	BF(0.704)--->1.455
IND40.2	8	3	AIM	0.793	AIF	0.145	3.704 --> 4.012

Appendix Table B-3. (hu)MANid (Berg and Kenyhercz 2017) Outputs using Measurements Taken by JHE and Linear Discriminant Analysis (LDA).

Case Number	# of Meas. Used	# of Pops. Used	Predicted Ancestry	Posterior Probability	Second Predicted Ancestry	Posterior Probability	Distances to Centroid
IND27	9	9	HISPM	0.541	BM	0.207	0.201 --> 0.647
IND32	3	3	HISPF	0.562	BF	0.375	1.558 --> 1.798
IND33	5	9	NEAM	0.272	HISPM	0.245	0.663 --> SEAM(0.856)
IND40.1	8	5	AIF	0.577	SEAF	0.152	BF(0.539) --> 0.948
IND40.2	8	8	AIM	0.677	AIF	0.188	0.458 --> BM(1.083)

Appendix Table B-4. (hu)MANid (Berg and Kenyhercz 2017) Outputs using Measurements Taken by JHE and Mixture Discriminant Analysis (MDA).

Case Number	# of Meas. Used	# of Pops. Used	Predicted Ancestry	Posterior Probability	Second Predicted Ancestry	Posterior Probability	Distances to Centroid
IND27	9	9	HISPM	0.334	BM	0.317	BM(0.420)--->0.716
IND32	3	3	BF	0.514	WM	0.423	HISPF(1.720) --> 1.891
IND33	5	9	HISPM	0.447	SEAM	0.18	NEAM(1.310)--->2.356
IND40.1	8	5	AIF	0.297	SEAF	0.282	0.297 --> 0.282
IND40.2	8	8	AIM	0.548	AIF	0.173	1.088 --> 1.777

Appendix Table B-5. (hu)MANid (Berg and Kenyhercz 2017) Outputs using Measurements Taken for Intra-observer Error and Linear Discriminant Analysis (LDA).

Case Number	# of Meas. Used	# of Pops. Used	Predicted Ancestry	Posterior Probability	Second Predicted Ancestry	Posterior Probability	Distances to Centroid
IND6	9	8	AIM	0.281	HISPM	0.228	2.287 --> NEAM(2.364)
IND11	9	3	AIM	0.388	HISPM	0.185	0.287 -- > BM(0.797)
IND27	9	8	HISPM	0.433	BM	0.228	0.804 --> 0.847
IND33	4	5	AIM	0.358	BM	0.199	0.376 --> 1.240
IND40.1	8	3	AIF	0.913	BF	0.071	1.210 --> 2.564
IND40.2	8	8	AIM	0.503	AIF	0.327	0.563 --> BM(0.970)
IND47	6	10	AIF	0.388	BF	0.187	0.108 --> 0.631
IND50.2	9	9	HISPM	0.356	WM	0.204	0.676 --> SEAF(0.703)

Appendix Table B-6. (hu)MANid (Berg and Kenyhercz 2017) Outputs using Measurements Taken for Intra-observer Error and Mixture Discriminant Analysis (MDA).

Case Number	# of Meas. Used	# of Pops. Used	Predicted Ancestry	Posterior Probability	Second Predicted Ancestry	Posterior Probability	Distances to Centroid
IND6	9	8	WM	0.325	AIM	0.219	NEAM(0.634)--->2.201
IND11	9	3	NEAM	0.231	HISPM	0.227	AIM(0.485) --> 1.285
IND27	9	8	HISPM	0.318	BM	0.305	SEAF(0.625)--->1.241
IND33	4	5	BM	0.253	HISPM	0.24	AIM(1.746)--->2.472
IND40.1	8	3	AIF	0.706	BF	0.254	1.016 --> 1.333
IND40.2	8	8	AIM	0.306	BM	0.234	1.078 --> AIF(1.860)
IND47	6	10	AIF	0.198	BF	0.182	0.272 --> 0.661
IND50.2	9	9	WM	0.702	HISPM	0.153	SEAF(1.362)--->1.838